

# Analytic Lagrangian tori for the planetary many-body problem\*

Luigi Chierchia

Fabio Pusateri

## Abstract

In 2004 J. Féjoz [7], completing investigations of M. Herman's [9], gave a complete proof of "Arnold's Theorem" [1] on the planetary many-body problem, establishing, in particular, the existence of a positive measure set of smooth ( $C^\infty$ ) Lagrangian invariant tori for the planetary many-body problem. Here, using Rüssmann's 2001 KAM theory [16], we prove the above result in the real-analytic class.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Analytic Lagrangian tori for properly degenerate systems</b>	<b>4</b>
2.1	Rüssmann's theorem for weakly nondegenerate systems . . . . .	5
2.2	A KAM Theorem for properly-degenerate systems . . . . .	7
2.3	Proof of Theorem 4 . . . . .	9
2.3.1	Step 1: Normal forms for properly-degenerate systems . . .	10
2.3.2	Step 2: Amounts of nondegeneracy of the normal form . . .	16
2.3.3	Conclusion of the Proof of Theorem 4 . . . . .	17
<b>3</b>	<b>Proof of Theorem 2</b>	<b>21</b>
	<b>Bibliography</b>	<b>25</b>

## 1 Introduction

The planetary many-body problem consists in studying the evolution of  $(1 + n)$  bodies (point masses), subject only to the mutual gravitational attraction, in the case one of bodies (the "Sun") has mass  $m_0$  considerably larger than the masses  $m_i$  of the remaining  $n$  bodies (the "planets"). The Newtonian evolution equations for such problem (in suitable units) are given by

$$m_j \ddot{q}_j = \sum_{k \neq j} m_j m_k \frac{q_k - q_j}{|q_k - q_j|^3}, \quad j = 0, 1, \dots, n, \quad (1)$$

---

\***Acknowledgments.** We are indebted with Jacques Féjoz for many enlightening discussions. This work was partially supported by the Italian MIUR project "Metodi variazionali e equazioni differenziali nonlineari".

where  $q_j = q_j(t) \in \mathbb{R}^3$  denotes the position at time  $t$  of the  $j^{\text{th}}$  body,  $|\cdot|$  denotes the euclidean norm and  $\dot{\cdot}$  denotes time derivative.

In [1, Chapter III, p. 125], V.I. Arnold made the following<sup>1</sup>:

**Arnold’s statement:** “In the  $n$ -body problem there exists a set of initial conditions having a positive Lebesgue measure and such that, if the initial positions and velocities belong to this set, the distances of the bodies from each other will remain perpetually bounded.”

As well known, such statement solves a fundamental problem considered, for several centuries, by astronomers and mathematicians. However, Arnold considered in details only the planar three-body case<sup>2</sup> and it appears that his indication for extending the result to the general case contains a flaw; compare end of §1.2, p. 1524 in [7].

A complete general proof of Arnold’s statement was given only in 2004, when J. Féjoz, completing the work of M. Herman, proved the following<sup>3</sup>

**Theorem 1 (Arnold, Herman, Féjoz [7, §1.2, p. 1523, THÉORÈME 1]).**

*Si le maximum  $\epsilon = \max\{m_j/m_0\}_{j=1,\dots,n}$  des masses des planètes rapportées à la masse du soleil est suffisamment petit, les équations (1) admettent, dans l’espace des phases au voisinage des mouvements képlériens circulaires et coplanaires, un ensemble de mesure de Lebesgue strictement positive de conditions initiales conduisant à des mouvements quasipériodiques.*

The beautiful proof of this result given in [7] (see also [8]) relies, on one side, on the elegant  $C^\infty$  KAM theory worked out by Herman (§ 2÷5 in [7]), and, on the other side, on the analytical celestial mechanics worked out, especially, by Poincaré and clarified and further investigated in Paris in the late 1980’s by A. Chenciner and J. Laskar in the *Bureau des Longitudes*<sup>4</sup> and later by Herman himself.

The invariant tori associated to the motions provided by Theorem 1, in view of the just mentioned KAM tools, are  $C^\infty$ . Now, since the many-body problem is formulated in terms of real-analytic functions, it appears somewhat more natural to seek for *real-analytic invariant manifolds*. This is the problem addressed in this paper. In particular, we shall give a new proof of Arnold’s statement, proving the following

**Theorem 2.** *If  $\epsilon = \max\{m_j/m_0\}_{j=1,\dots,n}$  is small enough, there exists a strictly positive measure set of initial conditions for the  $(1+n)$ -planetary problem (1), whose time evolutions lie on real-analytic Lagrangian tori in the  $6n$ -dimensional phase space*

$$\mathcal{M} := \left\{ (q, p) \in \mathbb{R}^{6(1+n)} : q_j \neq q_k, \forall j \neq k \text{ and } \sum_{j=0}^n p_j = 0 = \sum_{j=0}^n m_j q_j \right\},$$

*endowed with the restriction of the standard symplectic form  $\sum_{j=0}^n dq_j \wedge dp_j = \sum_{\substack{0 \leq j \leq n \\ 1 \leq k \leq 3}} dq_{j,k} \wedge dp_{j,k}$ .*

---

<sup>1</sup>The integer  $n$  in Arnold’s statement corresponds to the above  $(1+n)$ .

<sup>2</sup>Few lines after the above reported statement in [1, Chapter III, p. 125], Arnold says: “We shall consider only the plane three-body problem in detail. [...] In the final section a brief indication is given of the way in which the fundamental theorem of Chapter IV is applied in the investigation of the planetary motions in the plane and spatial many-body problems.”

<sup>3</sup>For a more detailed statement, see, footnote 6 below.

<sup>4</sup>Compare, e.g., the *Notes Scientifiques et techniques du Bureau des Longitudes* **S 026** and **S 028** by, respectively, Chenciner & Laskar and Chenciner.

**Remark 1.** Let us collect here a few observations concerning the above statements and respective proofs.

- (i) The proof of Theorem 2 given below is similar in strategy to that in [7] but technically different and it is based on an *analytic* (rather than smooth) KAM theory for properly-degenerate Hamiltonian systems (see, also, point (iv) below). On the other hand, it is conceivable – notwithstanding the presence of strong degeneracies (see point (ii) and (iii) below) – to prove regularity and uniqueness results for the planetary problem so as to deduce that the invariant tori in [7] are indeed analytic, a fact which does not follow from our proofs<sup>5</sup>.
- (ii) The evolution equations (1) are Hamiltonian and admit seven integrals, namely, the Hamiltonian (energy)  $H := \sum_{j=0}^n \frac{|p_j|^2}{2m_j} - \sum_{0 \leq k < j \leq n} \frac{m_j m_k}{|q_k - q_j|}$ , the three components of the total linear momentum  $M := \sum_{j=0}^n p_j$  and the three components of the total angular momentum  $C := \sum_{j=0}^n p_j \times q_j$ , where ‘ $\times$ ’ denotes the usual skew vector product in  $\mathbb{R}^3$ . As a reflection of the invariance of Newton’s equation (1) under changes of inertial reference frames, the Hamiltonian system associated to the  $(1+n)$ -body problem may be studied on the symplectic, invariant  $6n$ -dimensional manifold  $\mathcal{M}$  defined above, where, besides the total linear momentum, also the coordinates of the barycenter of the system vanish (“reduction of the total linear momentum”). However, the reduced  $(1+n)$ -body Hamiltonian still admits, besides the energy, three integrals given by the components of  $C = (C_x, C_y, C_z)$ . Incidentally, such integrals are not commuting since, if  $\{\cdot, \cdot\}$  denotes the natural Poisson bracket on  $\mathcal{M}$ , one has the cyclical relations  $\{C_x, C_y\} = C_z$ ,  $\{C_y, C_z\} = C_x$  and  $\{C_z, C_x\} = C_y$ ; but for example  $|C|^2$  and  $C_z$  are two independent, commuting integrals.
- (iii) The reasons why, notwithstanding the development of KAM theory in the early 1960’s, it took so long to give a complete proof of Arnold’s statement are technical in nature and are related to the strong degeneracies of the planetary problem (degeneracies, which are related to the abundance of integrals mentioned in the preceding item). The planetary  $(1+n)$ -body problem is perturbative, the unperturbed limit being obtained by considering  $n$  decoupled two-body problems formed by the Sun and the  $j^{\text{th}}$  planet. Now, the two-body problem in space is a three-degrees-of-freedom problem, but, once it is put into (Delaunay) action-angle variables, it depends only on one action (the action  $L$  proportional to the square root of the semi-major axis of the Keplerian ellipse on which the two bodies revolve). Systems of this kind are called *properly degenerate* and standard KAM theory does not apply. This difficulty, however, was overcome by Arnold – essentially by refined normal form theory – in the case of the planar three-body case, to which he could apply his “fundamental theorem” [1, chapter IV]. Indeed, Arnold’s approach, in view of Jacobi’s reduction of the nodes, could be extended [15] to the spatial three-body case ( $n = 2$ ) but not to the general case (spatial,  $n > 2$ ). Indeed, when  $n > 2$ , in space, a nice symplectic reduction (corresponding to the reduction of the nodes for  $n = 2$ ) is not yet known.

Furthermore (but not independently), in higher dimension, there appear two *secular resonances* (see Eq. (52) below), which prevent direct application of

---

<sup>5</sup>Recent interesting progresses in the study of uniqueness of invariant Lagrangian manifolds appeared in [4] and, especially, in [6]; however, as far as regularity is concerned, to the best of our knowledge, the only complete proven statement is Theorem 4, §4, p. 34 in [17], which covers only the  $C^\infty$  nondegenerate case.

any kind of KAM machinery. We mention that the way we overcome, here, this last difficulty is slightly different from that used in [7]: roughly speaking, in [7] it is introduced a modified Hamiltonian, which is then considered on the symplectic submanifold of vertical total angular momentum; here, we consider, instead, an extended phase space by adding an extra degree-of-freedom and consider on it a modified nondegenerate Hamiltonian.

- (iv) The main technical tool for us is the analytic KAM theory for weakly nondegenerate systems worked out by R  bmann in [16]; the main results of R  bmann’s theory (in the case of Lagrangian tori) are recalled in § 2.1 (see, also, Lemma 8 in § 2.3.3). The extension of this theory to properly degenerate systems is explained in § 2.2 and proved in § 2.3 (which constitutes the longest and most technical part of the paper). Finally, in § 3, using several results reported in [7], the proof of Theorem 2 is given.
- (v) Finally, we mention very briefly a few questions related to the context considered here:
  - Describe, in detail, the motions that take place on the Lagrangian tori. Let us clarify this point. From the proof of Theorem 2 given below, in view of the indirect argument used ([7, Lemma 82, p. 1578]) we cannot conclude that the ‘true’ motion is quasi-periodic; on the other hand, using different arguments, Arnold and F  joz say that the motion, in the general case, is quasi-periodic and takes place on  $(3n - 1)$ -dimensional tori<sup>6</sup>. Moreover, in the spatial three-body case ( $n = 2$ ) the Lagrangian tori are actually 4-dimensional (not  $5 = 3n - 1$ ) and the number of *independent* frequencies is 4 (compare [15]).
  - Find a ‘good’ set of analytic symplectic variables for the general spatial many-body problem.
  - Give asymptotic (as  $\epsilon \rightarrow 0$ ) estimates on the measure of Lagrangian invariant tori.
  - Apply some of the above result to a subsystem of the Solar system (for some progress in this direction, see [5]).

## 2 Analytic Lagrangian tori for properly degenerate systems

In this section we first recall a result due to R  bmann concerning analytic perturbations of weakly nondegenerate Hamiltonian systems (§ 2.1) and then show how such result may be used to give an analytic version of Herman’s  $C^\infty$  KAM theorem on

---

<sup>6</sup> Arnold ([1, p. 127]): “Thus, the Lagrangian motion is conditionally periodic and to the  $n_o$  ‘rapid’ frequencies of the Keplerian motion are added  $n_o$  (in the plane problem) or  $2n_o - 1$  (in the space problem) ‘slow’ frequencies of the secular motions”.

F  joz ([7, p. 1566]): **TH  OREME 60.** *Pour toute valeur des masses  $m_0, m_1, \dots, m_n > 0$  et des demi grands axes  $a_1 > \dots > a_n > 0$ , il existe un r  el  $\epsilon_0 > 0$  tel que, pour tout  $\epsilon$  tel que  $0 < \epsilon < \epsilon_0$ , le flot de l’hamiltonien  $F$  (d  fini en (28)) poss  de un ensemble de mesure de Lebesgue strictement positive de tores invariants de dimension  $3n - 1$ , de classe  $C^\infty$ , quasip  riodiques et  $\epsilon$ -proches en topologie  $C^0$  des tores k  pl  riens de demi grands axes  $(a_1, \dots, a_n)$  et d’excentricit  s et d’inclinaisons relatives nulles; de plus, quand tend vers z  ro la densit   des tores invariants au voisinage de ces tores k  pl  riens tend vers un.*

properly degenerate systems (i.e., nearly-integrable systems, which when the perturbation parameter vanishes depend on less action variables than the number of degrees of freedom). The statement of the analytic theorem for properly degenerate systems is given in § 2.2 and its proof in § 2.3.

## 2.1 Rüßmann's theorem for weakly nondegenerate systems

We start with fixing some notation.

- If  $a, b \in \mathbb{R}^n$  then  $\langle a, b \rangle := \sum_{i=1}^n a_i b_i$  and  $|a| := |a|_2 := \langle a, a \rangle^{\frac{1}{2}}$ ;
- if  $g$  is a  $\mu$ -times continuously differentiable function ( $\mu \in \mathbb{N}$ ) from an open set  $B \subset \mathbb{R}^n$  to  $\mathbb{R}^m$ , the  $\mu$ -th (tensor) derivative of  $g$  in  $b \in B$  is denoted by  $(a_1, \dots, a_\mu) \rightarrow \partial^\mu g(b)(a_1, \dots, a_\mu)$ ,  $a_j \in \mathbb{R}^n, j = 1, \dots, \mu$ ; if  $a_1 = \dots = a_\mu$ , we shall write  $\partial^\mu g(b)(a)^\mu$
- $|\partial^\mu g(b)| := \max_{a \in \mathbb{R}^n, |a|=1} |\partial^\mu g(b)(a, \dots, a)|$  and  $|\partial^\mu g|_A := \sup_{b \in A} |\partial^\mu g(b)|$ ;
- $C^\mu(B, \mathbb{R}^m)$  will denote the Banach space of all  $\mu$ -times continuously differentiable functions  $g : B \rightarrow \mathbb{R}^m$  with bounded derivatives up to order  $\mu$ , endowed with the norm  $|g|_B^\mu = \sup_{0 \leq \nu \leq \mu} |\partial^\nu g|_B < \infty$ .

The key notion of nondegeneracy is the following.

**Definition 1** (Rüßmann nondegeneracy condition<sup>7</sup>). *A real-analytic function*

$$\omega : y \in B \subset \mathbb{R}^n \longrightarrow \omega(y) = (\omega_1(y), \dots, \omega_m(y)) \in \mathbb{R}^m$$

*is called **R-nondegenerate** if  $B$  is a non-empty open connected set in  $\mathbb{R}^n$  and if for any  $c = (c_1, \dots, c_m) \in \mathbb{R}^m \setminus \{0\}$  one has*

$$y \longrightarrow \langle c, \omega \rangle := \sum_{i=1}^m c_i \omega_i \neq 0$$

*or equivalently if the range  $\omega(B)$  of  $\omega$  does not lie in any  $(m-1)$ -dimensional linear subspace of  $\mathbb{R}^m$ . We call  $\omega$  **R-degenerate** if it is not R-nondegenerate.*

The following lemma is a simple consequence of R-nondegeneracy and analyticity:

**Lemma 1.** *Let  $\omega : B \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$  be R-nondegenerate. Then for any non empty compact set  $\mathcal{K} \subset B$  there exist numbers  $\mu_0 = \mu_0(\omega, \mathcal{K}) \in \mathbb{Z}_+$  and  $\beta = \beta(\omega, \mathcal{K}) > 0$  such that*

$$\max_{0 \leq \mu \leq \mu_0} \left| \partial_y^\mu \langle c, \omega(y) \rangle^2 \right| \geq \beta, \quad \forall c \in \mathcal{S}^{m-1}, \quad \forall y \in \mathcal{K} \quad (2)$$

*where  $\mathcal{S}^{m-1} := \{c \in \mathbb{R}^m : |c|_2 = 1\}$ .*

For the proof see Lemma 18.2 on page 185 of [16].

In view of Lemma 1 one can give the following

---

<sup>7</sup>This terminology seems to be, nowadays, standard (see, e.g., [18]); however many authors, besides Rüßmann, contributed to its formulation; among them: Arnold, Margulis, Pyartli, Parasyuk, Bakhtin, Sprindzhuk and others.

**Definition 2.** Let  $\mathcal{K}$  and  $B$  as in the preceding lemma and let  $\omega : y \in B \longrightarrow \mathbb{R}^m$  be a real-analytic and  $R$ -nondegenerate function. We define  $\mu_0(\omega, \mathcal{K}) \in \mathbb{Z}_+$ , the **index of nondegeneracy** of  $\omega$  with respect to  $\mathcal{K}$ , as the smallest positive integer such that

$$\beta := \min_{y \in \mathcal{K}, c \in \mathcal{S}^{m-1}} \max_{0 \leq \mu \leq \mu_0} \left| \partial^\mu \langle c, \omega(y) \rangle^2 \right| > 0. \quad (3)$$

The number  $\beta = \beta(\omega, \mathcal{K})$  is called **amount of nondegeneracy** of  $\omega$  with respect to  $\mathcal{K}$ .

**Remark 2.** If a real-analytic function  $\omega : B \rightarrow \mathbb{R}^m$  admits the existence of  $\mu_0$  and  $\beta$  as in (3), for some compact  $\mathcal{K}$  containing an open ball, then it is  $R$ -nondegenerate.

The following result, which concerns the existence of maximal (Lagrangian) tori only, is a particular case of the main Theorem in [16], where also lower dimensional tori are treated<sup>8</sup>.

**Theorem 3 (Rüßmann, 2001).** Let  $\mathcal{Y}$  be an open connected set of  $\mathbb{R}^n$  and  $\mathbb{T}^n$  the usual  $n$ -dimensional torus  $\mathbb{R}^n/2\pi\mathbb{Z}^n$ . Consider a real-analytic Hamiltonian

$$H(x, y) = h(y) + P(x, y)$$

defined for  $(x, y) \in \mathbb{T}^n \times \mathcal{Y}$  endowed with the standard symplectic form  $dx \wedge dy$ . Let  $\mathcal{K}$  be any compact subset of  $\mathcal{Y}$  with positive  $n$ -dimensional Lebesgue measure  $\text{meas}_n \mathcal{K} > 0$  and fix  $0 < \epsilon^* < \text{meas}_n \mathcal{K}$ . Let  $\mathcal{A}$  be an open set in  $\mathbb{C}^n/2\pi\mathbb{Z}^n \times \mathbb{C}^n$  on which  $H$  can be analytically extended and such that  $\mathbb{T}^n \times \mathcal{K} \subset \mathcal{A}$ . Assume that the frequency application  $\omega := \nabla h$  is  $R$ -nondegenerate on  $\mathcal{Y}$ ; let  $\mu$  be any integer greater or equal to  $\mu_0(\omega, \mathcal{K})$  (the index of nondegeneracy of  $\omega$  with respect to  $\mathcal{K}$ ) and let  $\beta$  be as in (3) with  $\mu_0$  replaced by  $\mu$ .

Then, for any fixed  $\tau > n\mu$ , there exist  $\epsilon_0 = \epsilon_0(\epsilon^*, n, \mu, \beta, \tau, \omega, \mathcal{K}) > 0$  and  $\gamma = \gamma(\epsilon^*, n, \tau, \mu, \beta, \omega, \mathcal{K}) > 0$  such that if

$$|P|_{\mathcal{A}} := \sup_{\mathcal{A}} |P| \leq \epsilon_0 \quad (4)$$

the following is true. There exist a compact set

$$\mathcal{K}^* \subset \mathcal{K} \quad \text{with} \quad \text{meas}_n \mathcal{K}^* > \text{meas}_n \mathcal{K} - \epsilon^* \quad (5)$$

and a Lipschitz mapping

$$X : (b, \xi, \eta) \in \mathcal{K}^* \times \mathbb{T}^n \times \mathcal{U} \longrightarrow \mathbb{T}^n \times \mathcal{Y},$$

where  $\mathcal{U}$  is an open neighborhood of the origin in  $\mathbb{R}^n$ , such that:

(i) the mapping

$$(\xi, \eta) \longmapsto (x, y) = X(b, \xi, \eta)$$

defines, for every  $b \in \mathcal{K}^*$ , a real-analytic symplectic transformation<sup>9</sup> close to the identity on  $\mathbb{T}^n \times \mathcal{U}$ ;

<sup>8</sup>See Theorem 1.7 in [16, p. 127]. We refer to [13] for more details on how to obtain Theorem 3 from the general results in [16].

<sup>9</sup>I.e. it preserves the symplectic form  $dx \wedge dy$ .

(ii) the map

$$(b, \xi) \in \mathcal{K}^* \times \mathbb{T}^n \longrightarrow X_0(b, \xi) := X(b, \xi, 0)$$

is a bi-Lipschitz homeomorphism;

(iii) the transformed Hamiltonian  $H^* := H \circ X$  is in the form<sup>10</sup>:

$$H^*(b, \xi, \eta) = h^*(b) + \langle \omega^*(b), \eta \rangle + O(|\eta|^2)$$

for every  $b \in \mathcal{K}^*$  and  $(\xi, \eta) \in \mathbb{T}^n \times \mathcal{U}$ ;

(iv) the new frequency vector  $\omega^*$  satisfies for all  $b$  in  $\mathcal{K}^*$  the Diophantine inequality

$$|\langle k, \omega^*(b) \rangle| \geq \frac{\gamma}{|k|_2^\tau}, \quad \forall k \in \mathbb{Z}^n \setminus \{0\}. \quad (6)$$

**Remark 3.** (i) From Theorem 3 we immediately obtain that for any  $b \in \mathcal{K}^*$  the  $n$ -dimensional tori

$$\mathcal{T}_b := X_0(b, \mathbb{T}^n)$$

are invariant for  $H$  and the  $H$ -dynamics is analytically conjugate to  $\xi \rightarrow \xi + \omega^*(b)t$ . Furthermore, as it follows from (5) and point (ii), the measure of  $\cup_{b \in \mathcal{K}^*} \mathcal{T}_b$  is proportional to  $(\text{meas}_n \mathcal{K} - \epsilon^*)(2\pi)^n$  and, hence, tends to the full measure linearly when  $\epsilon^*$  tends to 0.

(ii) A (technical) difference between Rüssmann's Theorem and the formulation given above in Theorem 3 is the choice of  $\mu$  as any integer greater or equal than the actual index of nondegeneracy of  $\omega$ , while in [16]  $\mu$  is chosen equal to the index of nondegeneracy of  $\omega$ . In fact, it is easy to check<sup>11</sup> that Rüssmann's theorem holds in this slightly more general case, which will however be important in our applications.

Another difference of Theorem 3 above with respect to Rüssmann's original formulation, concerns the way the small divisors are controlled. Rüssmann uses a very general approach based upon "approximation functions"; however, such approach is too general for our application and cannot be applied directly. Nevertheless, it is easy to follow a more classical approach<sup>12</sup> based upon Diophantine inequalities of the form (6), which will be good enough for the application to properly degenerate systems; compare also remark 5, (ii) below.

## 2.2 A KAM Theorem for properly-degenerate systems

Let  $d$  and  $p$  be positive integers; let  $\mathcal{B}$  an open set in  $\mathbb{R}^d$ ,  $\mathcal{U}$  some open neighborhood of the origin in  $\mathbb{R}^{2p}$  and  $\epsilon$  a "small" real parameter. Consider a Hamiltonian function  $H_\epsilon$  of the form

$$H_\epsilon(\varphi, I, u, v) = h(I) + \epsilon f(\varphi, I, u, v), \quad (7)$$

real-analytic for

$$(\varphi, I, (u, v)) \in \mathbb{T}^d \times \mathcal{B} \times \mathcal{U} =: \mathcal{M}$$

<sup>10</sup>Here and in what follows  $f(\eta) = O(g(\eta))$  means that there exists a constant  $C$  such that  $|f(\eta)| \leq C|g(\eta)|$  for small enough  $\eta$ .

<sup>11</sup>See, e.g., [13].

<sup>12</sup>Compare, again, [13].

where  $\mathcal{M}$  is endowed with the standard symplectic form

$$d\varphi \wedge dI + du \wedge dv.$$

The “perturbation”  $f$  is assumed to have the form

$$\begin{cases} f(\varphi, I, u, v) = f_0(I, u, v) + f_1(\varphi, I, u, v), & \int_{\mathbb{T}^d} f_1(\varphi, I, u, v) d\varphi = 0 \\ f_0(I, u, v) = f_{00}(I) + \sum_{j=1}^p \Omega_j(I) \frac{u_j^2 + v_j^2}{2} + O(|(u, v)|^3). \end{cases} \quad (8)$$

Observe that the Hamiltonian  $h + \epsilon f_0$  possesses for every  $\bar{I} \in \mathcal{B}$  the invariant isotropic (non-Lagrangian) torus

$$\mathcal{T}_{\bar{I}}^d := \mathbb{T}^d \times \{\bar{I}\} \times \{0\} \subset \mathcal{M}$$

with corresponding quasi-periodic flow

$$\varphi(t) = \left( \partial_I h(\bar{I}) + \epsilon \partial_I f_{00}(\bar{I}) \right) t + \varphi_0 \quad I(t) \equiv \bar{I} \quad (u(t), v(t)) \equiv 0.$$

The purpose is to find Lagrangian invariant tori for  $H_\epsilon$  close to  $(d+p)$ -tori of the form

$$\mathcal{T}_{\bar{I}, w}^{d+p} = \mathbb{T}^d \times \{\bar{I}\} \times \{(u, v) \in \mathbb{R}^{2p}, |(u_j, v_j)|^2 = 2w_j, \forall j = 1, \dots, p\} \quad (9)$$

for  $\bar{I}$  in  $\mathcal{B}$  and  $w \in (\mathbb{R}_+)^p$  small.

**Theorem 4.** *Consider a real-analytic Hamiltonian function  $H_\epsilon$  as in (7) and (8), and assume that the “frequency map”*

$$I \in \mathcal{B} \longrightarrow (\omega(I), \Omega(I)) := (\nabla h(I), \Omega_1(I), \dots, \Omega_p(I)) \in \mathbb{R}^d \times \mathbb{R}^p \quad (10)$$

*is  $R$ -nondegenerate. Then, if  $\epsilon$  is sufficiently small, there exists a positive measure set of phase space points belonging to real-analytical, Lagrangian,  $H_\epsilon$ -invariant tori, which are close to  $\mathcal{T}_{\bar{I}, w}^{d+p}$  as in (9) with  $w_j = O(\epsilon)$ ; furthermore, the  $H_\epsilon$ -flow on such tori is quasi-periodic with Diophantine frequencies.*

**Remark 4.** (i) The above Theorem may be viewed as the real-analytic version for Lagrangian tori of the  $C^\infty$  KAM Theorem by M. Herman contained in [7] (see, in particular, Theorem 57, page 1559). Under stronger nondegeneracy assumptions the above theorem corresponds to the “Fundamental Theorem” in [1].

(ii) The word “properly-degenerate” refers to the fact that for  $\epsilon = 0$  the Hamiltonian  $H_0$  depends on  $d$  action variables, while the number of degrees of freedom is  $d+p > d$ . In particular, the tori constructed in Theorem 4, as  $\epsilon \rightarrow 0$ , degenerate into lower dimensional (non Lagrangian) tori  $\mathcal{T}_{\bar{I}}^d$ .

(iii) The natural symplectic variables for the KAM theory of the Hamiltonian  $H_\epsilon$  are  $(I, \varphi)$  and (rather than the cartesian variables  $(q, p)$ ) the symplectic action-angle variables  $(w, \zeta)$ , where  $w_j = \frac{u_j^2 + v_j^2}{2}$  for  $j = 1, \dots, p$  and  $\zeta_j$  is the angle of the circle<sup>13</sup>  $|w_j| = \text{const}$ . Indeed, Theorem 4 has, in terms of such variables,

---

<sup>13</sup>Compare Eq. (27) below, where  $w$  is related to  $\rho$  by  $w = \rho^0 + \rho$ .



a natural reformulation, which gives a deeper insight into the structure of the invariant tori<sup>14</sup>:

**Theorem 5.** *Let  $H_\epsilon(\varphi, I, \zeta, w) = h(I) + \epsilon f(\varphi, I, \zeta, w)$  be real-analytic for  $(\varphi, I, \zeta, w) \in \mathbb{T}^d \times \mathcal{B} \times \mathbb{T}^p \times \{w \in \mathbb{R}^p : 0 < |w_j| < r\} =: \mathcal{M}$  for some open set  $\mathcal{B} \subset \mathbb{R}^d$  and  $0 < r$ ;  $\mathcal{M}$  is endowed with the symplectic form  $d\varphi \wedge dI + d\zeta \wedge dw$ . The perturbation  $f$  is of the form  $f = f_0(I, \zeta, w) + f_1$  with  $f_1$  having vanishing  $\varphi$ -mean value over  $\mathbb{T}^d$ ; furthermore  $f_0$  has the form  $f_0 = f_{00}(I) + \langle \Omega(I), w \rangle + o(|w|)$ . Then, if the frequency map  $I \in \mathcal{B} \rightarrow (\omega, \Omega) := (\partial_I h(I), \Omega(I)) \in \mathbb{R}^d \times \mathbb{R}^p$  is  $R$ -nondegenerate, and  $\epsilon$  is small enough, there exists a positive measure set of phase space points belonging to real-analytical, Lagrangian,  $H_\epsilon$ -invariant tori, which have the following parametrization:*

$$\begin{cases} \varphi = \theta + \tilde{\varphi}(\theta, \psi) \\ I = \bar{I} + \tilde{I}(\theta, \psi) \\ \zeta = \psi + \tilde{\zeta}(\theta, \psi) \\ w = \bar{w} + \tilde{w}(\theta, \psi) \end{cases}$$

where  $\bar{w}$  is a constant vector of norm  $2\epsilon$  and  $\tilde{\varphi}, \tilde{I}, \tilde{\zeta}$  and  $\tilde{w}$  real-analytic functions for  $(\theta, \psi) \in \mathbb{T}^d \times \mathbb{T}^p$  (with range, respectively, in  $\mathbb{T}^d, \mathbb{R}^d, \mathbb{T}^p$  and  $\mathbb{R}^p$ ) with

$$\begin{cases} \tilde{I} = O\left(\epsilon (\log \epsilon^{-1})^{-(\tau_0+1)}\right) \\ \tilde{w} = O\left(\epsilon^{\frac{\nu+1}{2}}\right) \\ \tilde{\varphi}, \tilde{\zeta} = O(\epsilon) \end{cases}$$

for suitable  $\nu \geq 4$  and<sup>15</sup>  $\tau_0 \geq d+p$ . Moreover, if  $(\omega, \Omega)$  is the above frequency map, the  $H_\epsilon$ -flow on such invariant tori is conjugated to

$$(\theta, \psi) \longrightarrow (\theta + \tilde{\omega}t, \psi + \epsilon \tilde{\Omega}t)$$

for a suitable Diophantine vector  $(\tilde{\omega}, \tilde{\Omega})$  satisfying

$$|\tilde{\omega} - \omega|, |\tilde{\Omega} - \Omega| = O(\epsilon).$$

## 2.3 Proof of Theorem 4

First of all, let us introduce some notation and make quantitative the assumptions of Theorem 4.

---

<sup>14</sup>In reformulating Theorem 4 in terms of the variables  $(\zeta, w)$  we shall often use the same symbols used above. The Proof of Theorem 5 will not be explicitly given since it follows easily from the proof of Theorem 4.

<sup>15</sup>See equations (16) and (21) below.

- For  $\delta > 0$ ,  $d \in \mathbb{N}$ ,  $A \subset \mathbb{R}^d$  or  $\mathbb{C}^d$  we denote

$$\begin{aligned} B^d(x_0, \delta) &:= \{x \in \mathbb{R}^d : |x - x_0| < \delta\}, & (x_0 \in \mathbb{R}^d), \\ D^d(x_0, \delta) &:= \{x \in \mathbb{C}^d : |x - x_0| < \delta\}, & (x_0 \in \mathbb{C}^d), \\ \mathbb{T}_\delta^d &:= \{x \in \mathbb{C}^d : |\operatorname{Im} x_j| < \delta, \operatorname{Re} x_j \in \mathbb{T}, \forall j = 1 \dots d\} \\ A + \delta &:= \bigcup_{x \in A} D^d(x, \delta) \end{aligned} \quad (11)$$

- We may assume that  $H_\epsilon$  in (7) and (8) can be holomorphically extended for

$$(\varphi, I, (u, v)) \in \mathbb{T}_\sigma^d \times (\mathcal{B} + r_0) \times (\mathcal{U} + r_1) =: \mathcal{M}_\star. \quad (12)$$

In particular  $H_\epsilon$  is real-analytic on  $\mathbb{T}^d \times B^d(I_0, s) \times B^{2p}(0, r_1)$  for any  $I_0$  in  $\mathcal{B}$  and  $s < r_0$ . Moreover, we shall denote

$$M_0 := \sum_{k \in \mathbb{Z}^d} \left( \sup_{(\mathcal{B}+r_0) \times (\mathcal{U}+r_1)} |f_k(I, u, v)| \right) e^{|k|_1 \sigma} \quad (13)$$

as the “sup-Fourier” norm of  $f$  and let

$$M_1 := \sup_{I \in \mathcal{B}+r_0} |(\omega(I), \Omega(I))|. \quad (14)$$

The proof of Theorem 4 is based upon **two preliminary steps**:

- 1 computation of a suitable normal form for  $H_\epsilon$ ;
- 2 quantitative estimates on the amount of the nondegeneracy of the normal form.

### 2.3.1 Step 1: Normal forms for properly-degenerate systems

**Proposition 1.** *Fix an integer  $\nu \geq 4$ . Then, there exists  $m > d$  (depending on  $\omega$ ), and, for  $\epsilon$  small enough, a point  $I_0 \in \mathcal{B}$  and a real-analytic canonical transformation<sup>16</sup>  $\Phi_\epsilon$  such that the following holds. Let*

$$s := O\left((\log \epsilon^{-1})^{-m}\right) \quad (15)$$

then  $B^d(I_0, s) \subset \mathcal{B}$  and  $\Phi_\epsilon : (\vartheta, r, \zeta, \rho) \longrightarrow (\varphi, I, u, v)$  satisfies

$$\Phi_\epsilon : \mathbb{T}^d \times B^d(0, s/5) \times \mathbb{T}^p \times B^p(0, \epsilon) \longrightarrow \mathbb{T}^d \times B^d(I_0, s) \times \mathcal{U}$$

and  $\hat{H}_\epsilon := H_\epsilon \circ \Phi_\epsilon$  takes the form

$$\hat{H}_\epsilon(\vartheta, r, \zeta, \rho) = N_\epsilon(r, \rho; \rho^0) + \epsilon^\nu P_\epsilon(\vartheta, r, \zeta, \rho; \rho^0) \quad (16)$$

with

$$N_\epsilon := \frac{1}{\epsilon} h(I_0 + \epsilon r) + \hat{g}(I_0 + \epsilon r) + \frac{1}{2} \hat{\Omega}(I_0 + \epsilon r) \cdot (\rho^0 + \epsilon \rho) + Q_{\epsilon, I_0 + \epsilon r}(\rho^0 + \epsilon \rho) \quad (17)$$

and:  $\rho^0$  in  $(\mathbb{R}_+)^p$  is some point having euclidean norm  $2\epsilon$ ;  $Q_{\epsilon, I_0 + \epsilon r}$  is a polynomial of degree  $\nu - 1$  starting with cubic terms;  $\hat{g}$ ,  $\hat{\Omega}$  and  $P_\epsilon$  are real-analytic functions. Furthermore, one has

$$\sup_{r \in B^d(0, s/5)} |\hat{\Omega}(I_0 + \epsilon r) - \Omega(I_0 + \epsilon r)| = O\left(\epsilon (\log \epsilon^{-1})^{2m-1}\right).$$

---

<sup>16</sup>Symplectic up to rescalings.

**Proof of Proposition 1.** We start by recalling a measure theoretical result due to Pyartli (see [14] or [16, Theorem 17.1]):

**Lemma 2 (Pyartli).** *Let  $\mathcal{K} \subset \mathbb{R}^d$  be a compact set; let  $\theta \in (0, 1)$  define  $\mathcal{K}_\theta := \cup_{y \in \mathcal{K}} B^d(y, \theta)$ . Let  $g : \mathcal{K}_\theta \rightarrow \mathbb{R}$  be a real-analytic function satisfying*

$$\min_{y \in \mathcal{K}} \max_{0 \leq \nu \leq \mu_0} |\partial^\nu g(y)| \geq \beta$$

*for some  $\beta > 0$ . Then there exists  $C = C(\mu_0, \beta, d, \mathcal{K}, \theta)$  such that*

$$\text{meas}_d \{y \in \mathcal{K} : |g(y)| \leq t\} \leq C |g|_{\mathcal{K}_\theta}^{\mu_0+1} t^{\frac{1}{\mu_0}}$$

*for any  $0 \leq t \leq \frac{\beta}{2\mu_0+2}$ .*

Pyartli's Lemma implies:

**Lemma 3.** *Let  $\mathcal{K}$  be a compact set with positive  $d$ -dimensional Lebesgue measure and let  $0 < \varepsilon^* < \text{meas}_d \mathcal{K}$ ; let  $\omega : \mathcal{K}_\theta \rightarrow \mathbb{R}^d$  be  $R$ -nondegenerate and let  $\mu_0$  and  $\beta$  be its index and amount of nondegeneracy with respect to  $\mathcal{K}$ . Let us denote by  $\mathcal{D}_{\gamma_0, \tau_0}^d$  the set of Diophantine vectors in  $\mathbb{R}^d$  with Diophantine constants  $\gamma_0, \tau_0$ , i.e. the set*

$$\mathcal{D}_{\gamma_0, \tau_0}^d := \left\{ \omega \in \mathbb{R}^d : |\langle \omega, k \rangle| \geq \frac{\gamma_0}{|k|^{\tau_0}}, \forall k \in \mathbb{Z} \setminus \{0\} \right\}.$$

*Then, if  $\gamma_0$  is sufficiently small and  $\tau_0 \geq d\mu_0$  one has*

$$\text{meas}_d \left( \mathcal{K} \cap \mathcal{D}_{\gamma_0, \tau_0}^d \right) \geq \text{meas}_d \mathcal{K} - \varepsilon^*. \quad (18)$$

**Proof of Lemma 3.** First of all observe that for any  $m \in \mathbb{Z}_+$ ,  $a \in \mathbb{R}^d$ ,  $k \in \mathbb{Z}^d \setminus \{0\}$  and  $b \in \mathcal{K}_\theta$

$$\left| \partial^m \langle \omega(b), k |k|^{-1} \rangle (a^m) \right| = \left| \langle \partial^m \omega(b)(a^m), k |k|^{-1} \rangle \right| \leq |\partial^m \omega(b)(a^m)|;$$

taking the sup over  $|a| = 1$ ,  $b \in \mathcal{K}_\theta$  and  $0 \leq m \leq \mu$  we get

$$\left| \langle \omega, k |k|^{-1} \rangle \right|_{\mathcal{K}_\theta}^\mu \leq |\omega|_{\mathcal{K}_\theta}^\mu < \infty$$

for any  $\mu \in \mathbb{Z}_+$ . Now we use this last inequality and Theorem 2, assuming  $\gamma_0 \leq \frac{\beta}{2\mu_0+2}$ , to estimate

$$\begin{aligned} \text{meas}_d \left( \mathcal{K} \setminus \mathcal{D}_{\gamma_0, \tau_0}^d \right) &= \text{meas}_d \bigcup_{k \in \mathbb{Z}^d \setminus \{0\}} \left\{ b \in \mathcal{K} : |\langle \omega(b), k \rangle| < \frac{\gamma_0}{|k|^{\tau_0}} \right\} \\ &\leq \sum_{k \in \mathbb{Z}^d \setminus \{0\}} \text{meas}_d \left\{ b \in \mathcal{K} : \left| \langle \omega(b), \frac{k}{|k|} \rangle \right| < \frac{\gamma_0}{|k|^{\tau_0+1}} \right\} \\ &\leq C(\mu_0, \beta, d, \mathcal{K}, \theta) |\omega|_{\mathcal{K}_\theta}^{\mu_0+1} \gamma_0^{\frac{1}{\mu_0}} \sum_{k \in \mathbb{Z}^d \setminus \{0\}} \frac{1}{|k|^{\frac{\tau_0+1}{\mu_0}}}. \end{aligned}$$

Since  $\tau_0 \geq d\mu_0$  this last sum converges and one has

$$\text{meas}_d \left( \mathcal{K} \setminus \mathcal{D}_{\gamma_0, \tau_0}^d \right) \leq \bar{C} \gamma_0^{\frac{1}{\mu_0}}$$

for a suitable  $\bar{C} = \bar{C}(\mu_0, \beta, d, \mathcal{K}, \theta, \omega, \tau_0)$ . Choosing  $\gamma_0 \leq (\bar{C}^{-1} \epsilon^\star)^{\mu_0}$  we obtain estimate (18).  $\blacksquare$

Now consider the real-analytic Hamiltonian  $H_\epsilon$  in (7) and (8). Let  $\nu_1, \nu_2 \geq 4$  be two integers to be later determined and set

$$K_1 := \frac{6}{\sigma} (\nu_1 - 1) \log \frac{1}{\epsilon M_0} \quad (19)$$

where  $M_0$  is defined by (13). Lemma 3 and the R-nondegeneracy of  $\omega$  assure the existence of  $I_0 \in \mathcal{B}$  such that  $\omega(I_0)$  belongs to  $\mathcal{D}_{\gamma_0, \tau_0}^d$  (for suitable  $\gamma_0$  and  $\tau_0$ ). Then, from Taylor's formula it follows that<sup>17</sup>

$$|\omega(I) \cdot k| \geq \alpha_1 > 0, \quad \forall k \in \mathbb{Z}^d, \quad 0 < |k|_1 \leq K_1, \quad \forall I \in D^d(I_0, s) \quad (20)$$

with<sup>18</sup>

$$s := O \left( (\log \epsilon^{-1})^{-(\tau_0+1)} \right) \quad \text{and} \quad \alpha_1 := O \left( (\log \epsilon^{-1})^{-\tau_0} \right). \quad (21)$$

Furthermore, we can assume that there exists  $\alpha_2$  (independent of  $\epsilon$ ) such that

$$|\Omega(I) \cdot k| \geq \alpha_2 > 0, \quad \forall k \in \mathbb{Z}^p, \quad 0 < |k|_1 \leq \nu_2, \quad \forall I \in D^d(I_0, s). \quad (22)$$

Next, we want to average  $H_\epsilon$  over the “fast angles”  $\varphi$  up to order  $\nu_1$ . To do this we shall apply the following classical “averaging lemma”, whose proof can be found in [3, Appendix A, p. 110].

**Lemma 4 (Averaging Lemma<sup>19</sup>).** *Let  $H_\epsilon, M_0, \sigma, \alpha_1$  and  $s$  be as above. Assume (20) holds with  $K_1$  as in (19). Then, if  $\epsilon$  is small enough, there exists a real-analytic symplectic transformation  $\Phi_\epsilon^1 : (\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v}) \rightarrow (\varphi, I, u, v)$  mapping*

$$\mathcal{M}_1 := \mathbb{T}_{\frac{\sigma}{6}}^d \times D^d(I_0, \frac{s}{2}) \times D^{2p}(0, \frac{r_1}{2}) \xrightarrow{\Phi_\epsilon^1} \mathcal{M}_0 := \mathbb{T}_\sigma^d \times D^d(I_0, s) \times D^{2p}(0, r_1)$$

that casts  $H_\epsilon$  into the Hamiltonian

$$H_\epsilon^1 := H_\epsilon \circ \Phi_\epsilon^1 = h + \epsilon f_0 + \tilde{g} + \tilde{f}$$

where  $\tilde{g} = \tilde{g}(\tilde{I}, \tilde{u}, \tilde{v})$  and  $\tilde{f}$  satisfy

$$\sup_{\mathcal{M}_1} |\tilde{g}| \leq C \frac{(\epsilon M_0)^2}{s \alpha_1}, \quad \sup_{\mathcal{M}_1} |\tilde{f}| \leq (\epsilon M_0)^{\nu_1} \quad (23)$$

for a suitable  $C = C(\sigma, \nu_1)$ .

<sup>17</sup>For  $k = (k_1, k_2, \dots, k_d) \in \mathbb{Z}^d$  we denote  $|k|_1 := \sum_{i=1}^d |k_i|$ ; recall also the definition of complex balls  $D^d$  in (11).

<sup>18</sup>This means that we can take  $m = \tau_0 + 1$  in (15).

<sup>19</sup>Lemma 4 can be immediately derived from Proposition A.1 in [3] with the following correspondences:  $\alpha_1 = \alpha$  for  $\alpha_1$  as in (20),  $K_1 = K$  for  $K_1$  as in (19),  $\epsilon M_0 = \epsilon$  for  $M_0$  as in (13),  $s = r, d$  for  $s$  as in (15) and (20),  $\{0\} = \Lambda$  and  $\epsilon f(I, \varphi, u, v)$  in (7) is just  $f(u, \varphi)$  in [3]; as a result one has that  $\epsilon f_0 + \tilde{g}$  and  $\tilde{f}$  are respectively given by  $g$  and  $f_\star$  in [3] with estimates (23) holding in view of the previous correspondences.

Thus, if we set  $\tilde{g} =: \epsilon^2 \bar{g}$  and  $\tilde{f} =: \epsilon^{\nu_1} \bar{f}$ , using (8) and (23) we have

$$\begin{cases} H_\epsilon^1(\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v}) = h(\tilde{I}) + \epsilon \left[ f_0(\tilde{I}, \tilde{u}, \tilde{v}) + \epsilon \bar{g}(\tilde{I}, \tilde{u}, \tilde{v}) \right] + \epsilon^{\nu_1} \tilde{f}(\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v}) \\ f_0(\tilde{I}, \tilde{u}, \tilde{v}) = f_{00}(\tilde{I}) + \sum_{j=1}^p \Omega_j(\tilde{I}) \frac{\tilde{u}_j^2 + \tilde{v}_j^2}{2} + O(|\tilde{u}, \tilde{v}|^3; \tilde{I}) . \end{cases} \quad (24)$$

From equation (24) we see that the application of averaging theory may cause, in general, a shift of order  $\epsilon$  of the elliptic equilibrium, which, before, was in the origin of  $\mathbb{R}^{2p}$ . Therefore, we focus our attention on the Hamiltonian function  $f_0 + \epsilon \bar{g}$  with the aim to find a real-analytic symplectic transformation restoring the equilibrium in the origin. An application of the standard Implicit Function Theorem yields the following<sup>20</sup>:

**Lemma 5.** *Let  $\mathcal{M}_2 := \mathbb{T}_{\frac{\epsilon}{\tau}}^d \times D^d(I_0, s/4) \times D^{2p}(0, r_1/4)$ ; then, provided  $\epsilon$  is sufficiently small, there exists a (close to the identity) real-analytic symplectic transformation*

$$\Phi_\epsilon^2 : (x, y, p, q) \in \mathcal{M}_2 \longrightarrow (\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v}) \in \mathcal{M}_1$$

such that  $H_\epsilon^2 := H_\epsilon^1 \circ \Phi_\epsilon^2$  is of the form

$$H_\epsilon^2(x, y, p, q) = h(y) + \epsilon \hat{g}(y, p, q) + \epsilon^{\nu_1} \hat{f}(x, y, p, q)$$

with  $\partial_p \hat{g}(y, 0, 0) = 0 = \partial_q \hat{g}(y, 0, 0)$ ,  $\hat{g}$  and  $\hat{f}$  real-analytic on  $\mathcal{M}_2$ .

Now, we need to control the frequencies associated to the modified Hamiltonian  $\hat{g}(y, 0, 0)$ :

**Lemma 6.** *If  $\epsilon$  is small enough then the eigenvalues of the Hamiltonian  $\hat{g}$ , i.e. the eigenvalues of  $\partial_{(p,q)}^2 J_{2p} \partial_{(p,q)}^2 \hat{g}(y, 0, 0)$ , are given by  $2p$  purely imaginary functions  $\pm i\hat{\Omega}_1, \dots, \pm i\hat{\Omega}_p$  verifying*

$$\sup_{y \in D^d(I_0, s/4)} |\hat{\Omega}(y) - \Omega(y)| = O\left(\epsilon (\log \epsilon^{-1})^{2\tau_0+1}\right) \quad (25)$$

for  $\tau_0$  as in (21).

**Proof of Lemma 6.** Consider the quadratic part of  $\hat{g}$ , that is the real-analytic  $2p \times 2p$  symmetric matrix  $\hat{A}(y) := \partial_{(p,q)}^2 \hat{g}(y, 0, 0)$ . Using the construction of  $\Phi_\epsilon^2$  in Lemma 5,  $\hat{g} = (f_0 + \epsilon \bar{g}) \circ \Phi_\epsilon^2$ , estimate (23) together with  $\tilde{g} = \epsilon^2 \bar{g}$  and the definition of  $s$  and  $\alpha_1$  in (21), equation (24) for  $f_0$  and Cauchy's estimate for derivatives of analytic functions, one has

$$\hat{A}(y) = \text{diag}(\Omega_1(y), \dots, \Omega_p(y), \Omega_1(y), \dots, \Omega_p(y)) + O\left(\epsilon (\log \epsilon^{-1})^{2\tau_0+1}\right) .$$

<sup>20</sup>Since  $\Omega_j \neq 0$  for every  $j$  in view of (22), we can apply the Implicit function Theorem to obtain, for small enough  $\epsilon$ , the existence of two functions  $u_0 = u_0(\tilde{I}, \epsilon)$  and  $v_0 = v_0(\tilde{I}, \epsilon)$  which are real-analytic for  $\tilde{I} \in D^d(I_0, s/4)$  and such that  $\nabla_{\tilde{u}, \tilde{v}}(f_0 + \epsilon \bar{g})(\tilde{I}, u_0, v_0) = 0$ . Furthermore, using (23) together with  $\tilde{g} = \epsilon^2 \bar{g}$  and (21), one has  $u_0, v_0 = O(\epsilon (\log \epsilon^{-1})^{2\tau_0+1})$ . The symplectic transformation in Lemma 5 is then generated by  $x \cdot \tilde{\varphi} + (p + u_0(x, \epsilon)) \cdot (\tilde{v} - v_0(x, \epsilon))$ .

<sup>21</sup> $J_{2p}$  denotes the standard  $2p \times 2p$  symplectic matrix.

Since  $\Omega_j \neq \Omega_k$  for<sup>22</sup>  $j \neq k$ , an application of the Implicit Function Theorem tells us that the eigenvalues of  $\hat{g}$  (that a priori might have non-zero real part) are  $O\left(\epsilon(\log \epsilon^{-1})^{2\tau_0+1}\right)$  close to  $\pm i\Omega_j$ . Now, as it is well known, eigenvalues of Hamiltonians always appear in quadruplets  $\pm\lambda, \pm\bar{\lambda}$ ; thus, from the simplicity of the eigenvalues of  $\hat{g}$  (holding for  $\epsilon$  small enough) one has that its eigenvalues are purely imaginary as claimed.  $\blacksquare$

By normal form theory (see corollary 8.7 of [2]) we can find a real-analytic symplectic transformation  $O(\epsilon)$ -close to the identity

$$\Phi_\epsilon^3 : (\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) \in \mathcal{M}_3 := \mathbb{T}_{\frac{s}{8}}^d \times D^d(I_0, \frac{s}{5}) \times D^{2p}(0, \frac{r_1}{5}) \longrightarrow (x, y, p, q) \in \mathcal{M}_2$$

with  $\tilde{y} = y$  and such that the transformed Hamiltonian function  $H_\epsilon^3 := H_\epsilon^2 \circ \Phi_\epsilon^3$ , which is real-analytic on  $\mathcal{M}_3$ , has the form

$$\begin{aligned} H_\epsilon^3(\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) &= h(\tilde{y}) + \epsilon \hat{g}_0(\tilde{y}) + \frac{\epsilon}{2} \sum_{j=1}^p \hat{\Omega}_j(\tilde{y}) (\tilde{p}_j^2 + \tilde{q}_j^2) \\ &\quad + \epsilon \tilde{g}_3(\tilde{y}, \tilde{p}, \tilde{q}) + \epsilon^{\nu_1} \tilde{f}_3(\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) \end{aligned}$$

where  $\hat{g}_0 := \hat{g}(\tilde{y}, 0, 0)$ ,  $\tilde{f}_3 := \tilde{f} \circ \Phi_\epsilon^3$  and  $\tilde{g}_3 := \hat{g}_3 \circ \Phi_\epsilon^3$  verifies

$$\sup_{\tilde{y} \in D^d(I_0, s/5)} |\tilde{g}_3(\tilde{y}, \tilde{p}, \tilde{q})| \leq C |\tilde{p}, \tilde{q}|^3 \quad \forall (\tilde{p}, \tilde{q}) \in D^{2p}(0, r_1/5).$$

Now let  $\tilde{g}_2(\tilde{y}, \tilde{p}, \tilde{q}) := \frac{1}{2} \sum_{i=1}^p \hat{\Omega}_i(\tilde{y}) (\tilde{p}_i^2 + \tilde{q}_i^2)$ , we want to put  $\tilde{g}_2 + \epsilon \tilde{g}_3$  into Birkhoff's normal form up to order  $\nu_2$ . In view of inequalities (22) and (25), provided  $\epsilon$  is small enough, we have

$$|\hat{\Omega}(\tilde{y}) \cdot k| \geq \frac{\alpha_2}{2}, \quad \forall k \in \mathbb{Z}^p, 0 < |k|_1 \leq \nu_2, \quad \forall \tilde{y} \in D^d(I_0, s/5). \quad (26)$$

By Birkhoff's normal form theory<sup>23</sup>, one obtains easily the following

**Lemma 7.** *If inequality (26) is satisfied, then there exist  $0 < r_\star < r'_1 \leq r_1/5$  and a real-analytic symplectic diffeomorphism  $\Phi_\epsilon^4 : (\theta, r, u, v) \rightarrow (\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q})$  mapping*

$$\mathcal{M}_4 := \mathbb{T}_{\frac{s}{8}}^d \times D^d(I_0, \frac{s}{5}) \times D^{2p}(0, r_\star) \xrightarrow{\Phi_\epsilon^4} \mathcal{M}_3' := \mathbb{T}_{\frac{s}{8}}^d \times D^d(I_0, \frac{s}{5}) \times D^{2p}(0, r'_1)$$

*leaving the origin and the quadratic part of  $H_\epsilon^3$  invariant, such that  $(\theta, r) = (\tilde{x}, \tilde{y})$  and  $H_\epsilon^4 := H_\epsilon^3 \circ \Phi_\epsilon^4$  is of the form*

$$\begin{aligned} H_\epsilon^4(\theta, r, u, v) &= h(r) + \epsilon \hat{g}_0(r) + \frac{\epsilon}{2} \sum_{j=1}^p \hat{\Omega}_j(r) (u_j^2 + v_j^2) + \\ &\quad + \epsilon Q_\star(r, u, v) + \epsilon R_\star(r, u, v) + \epsilon^{\nu_1} \tilde{f}_4(\theta, r, u, v) \end{aligned}$$

where:

---

<sup>22</sup>From (22).

<sup>23</sup>See, e.g., [10, Theorem 11, p. 43] or [13, section 3.4] for a quantitative version.

- $Q_\star$  is a polynomial of degree  $\lceil \frac{\nu_2}{2} \rceil$  in the variables  $I = (I_1, \dots, I_p)$  having the form

$$\langle \hat{\Omega}(r), I \rangle + \frac{1}{2} \langle T(r)I, I \rangle + \dots \quad \text{with} \quad I_j := \frac{1}{2}(u_j^2 + v_j^2)$$

with  $T(r)$  a  $2p \times 2p$  real-analytic matrix;

- $R_\star$  is a real-analytic function verifying  $|R_\star(r, u, v)| \leq C|(u, v)|^{\nu_2+1}$  for every  $(u, v) \in D^{2p}(0, r_\star)$  and  $r \in D^d(I_0, s/5)$ ;
- $\tilde{f}_4 := \tilde{f}_3 \circ \Phi_\epsilon^4$  is real-analytic on  $\mathcal{M}_4$ .

We may conclude the proof of Proposition 1. Following [7, pp. 1561–1562], we pass to symplectic polar coordinates in order to move  $R_\star$  to the perturbation of  $H_\epsilon^4$  with the help of a rescaling by a factor  $\epsilon$ . Let  $\rho^0 = (\rho_1^0, \dots, \rho_p^0)$  in  $(\mathbb{R}_+)^p$  be sufficiently close to the origin; consider, for a suitable  $\sigma_\star > 0$ , the real-analytic symplectic transformation  $\Phi_\epsilon^5 : (\theta, r, \zeta, \rho) \rightarrow (\theta, I_0 + r, z)$  mapping

$$\mathcal{M}_5 := \mathbb{T}_{\frac{\sigma}{8}}^d \times D^d(0, \frac{s}{5}) \times \mathbb{T}_{\sigma_\star}^p \times D^p(0, \frac{|\rho^0|}{2}) \xrightarrow{\Phi_\epsilon^5} \mathcal{M}_4$$

where

$$z_j = u_j + iv_j := \sqrt{2(\rho_j^0 + \rho_j)} e^{-i\zeta_j}. \quad (27)$$

The transformed Hamiltonian function  $H_\epsilon^5 := H_\epsilon^4 \circ \Phi_\epsilon^5$ , real-analytic on  $\mathcal{M}_5$ , assumes the form

$$\begin{aligned} H_\epsilon^5(\theta, r, \zeta, \rho) &= h(I_0 + r) + \epsilon \hat{g}_0(I_0 + r) + \frac{\epsilon}{2} \sum_{j=1}^p \hat{\Omega}_j(I_0 + r)(\rho_j^0 + \rho_j) \\ &\quad + \epsilon Q_{I_0+r}(\rho^0 + \rho) + \epsilon R(I_0 + r, \zeta, \rho^0 + \rho) + \epsilon^{\nu_1} \tilde{f}_5(\theta, r, \zeta, \rho; \rho^0) \end{aligned}$$

where

- $Q_{I_0+r} := Q_\star \circ \Phi_\epsilon^5$  is a polynomial of degree  $\lceil \frac{\nu_2}{2} \rceil$  with respect to  $\rho^0 + \rho$ , depending also on  $I_0 + r$ ;
- $R := R_\star \circ \Phi_\epsilon^5$  verifies

$$|R(I_0 + r, \zeta, \rho^0 + \rho)| \leq C|\rho^0|^{\frac{\nu_2+1}{2}}$$

for every  $\rho \in D^{2p}(0, |\rho^0|/2)$ ,  $r \in D^d(0, s/5)$  and  $\zeta \in \mathbb{T}_{\sigma_\star}^p$ ;

- $\tilde{f}_5 := \tilde{f}_4 \circ \Phi_\epsilon^5$  is real-analytic on  $\mathcal{M}_5$ .

Now, let  $A_\epsilon$  be the homothety given by

$$A_\epsilon : (\theta, r, \zeta, \rho) \longrightarrow (\theta, \epsilon r, \zeta, \epsilon \rho).$$

Even though  $A_\epsilon$  is not a symplectic map it preserves the structure of Hamilton's equations if we consider the Hamiltonian function  $H_\epsilon^6 := \frac{1}{\epsilon} H_\epsilon^5 \circ A_\epsilon$ . Explicitly we have

$$\begin{aligned} H_\epsilon^6(\theta, r, \zeta, \rho) &= \frac{1}{\epsilon} h(I_0 + \epsilon r) + \hat{g}_0(I_0 + \epsilon r) + \frac{1}{2} \hat{\Omega}(I_0 + \epsilon r) \cdot (\rho^0 + \epsilon \rho) \\ &\quad + Q_{\epsilon, I_0+\epsilon r}(\rho^0 + \epsilon \rho) + R(I_0 + \epsilon r, \epsilon \rho, \zeta; \rho^0) \\ &\quad + \epsilon^{\nu_1-1} \tilde{f}_6(\theta, r, \zeta, \rho; \rho^0). \end{aligned} \quad (28)$$

where  $\tilde{f}_6 := \tilde{f}_5 \circ A_\epsilon$ . Now we fix  $\rho^0 \in (\mathbb{R}_+)^p$  with  $|\rho^0| = 2\epsilon$  so that  $|R| \leq C\epsilon^{\frac{\nu_2+1}{2}}$ . Thus, if we choose  $\nu_1$  and  $\nu_2$  so that

$$\nu_1 - 1 = \left\lceil \frac{\nu_2 + 1}{2} \right\rceil := \nu \quad (29)$$

we may write

$$R(I_0 + \epsilon r, \epsilon \rho, \zeta; \rho^0) + \epsilon^{\nu_1-1} \tilde{f}_6(\theta, r, \zeta, \rho) =: \epsilon^\nu P_\epsilon(\theta, r, \zeta, \rho) \quad (30)$$

for a suitable function  $P_\epsilon$  real-analytic on  $\mathbb{T}_{\frac{d}{8}}^d \times D^d(0, s/5) \times \mathbb{T}_{\sigma_*}^p \times D^p(0, \epsilon)$ .

We have proved Proposition 1 with  $\hat{H}_\epsilon = H_\epsilon^6$  as in (28), (30).  $\blacksquare$

### 2.3.2 Step 2: Amounts of nondegeneracy of the normal form

**Proposition 2.** *Let  $N_\epsilon$  be as in (17). If  $\epsilon$  is small enough, the frequency map*

$$\hat{\Psi}_\epsilon : (r, \rho) \in B^d(0, s/5) \times B^p(0, \epsilon) \longrightarrow \left( \frac{\partial}{\partial r} N_\epsilon, \frac{\partial}{\partial \rho} N_\epsilon \right)$$

*is  $R$ -nondegenerate.*

*Moreover, let  $\bar{\mu}$  and  $\bar{\beta}$  denote respectively the index and the amount of nondegeneracy of the unperturbed frequency map (10) with respect to a closed ball  $\bar{B}^d(I_0, t) \subset \mathcal{B}$ , for some  $t > 0$  independent of  $\epsilon$ . Then, if we define  $\mathcal{K}_\epsilon := \bar{B}^d(0, s/10) \times \bar{B}^p(0, \epsilon/2)$  and let  $\hat{\mu}_\epsilon$  denote the index of nondegeneracy of  $\hat{\Psi}_\epsilon$  with respect to  $\mathcal{K}_\epsilon$  and*

$$\hat{\beta}_\epsilon := \min_{c \in \mathcal{S}^{d+p-1}} \min_{(r, \rho) \in \mathcal{K}_\epsilon} \max_{0 \leq \mu \leq \bar{\mu}} \left| \partial_{(r, \rho)}^\mu |\langle c, \hat{\Psi}_\epsilon \rangle|^2 \right|,$$

*one has*

$$\hat{\mu}_\epsilon \leq \bar{\mu} \quad \text{and} \quad \hat{\beta}_\epsilon \geq \frac{\epsilon^{\bar{\mu}+2} \bar{\beta}}{8}. \quad (31)$$

**Proof of Proposition 2.** From (17), it follows that the frequency map of  $N_\epsilon$  is given by

$$\hat{\Psi}_\epsilon(r, \rho) = \left( \omega(I_0 + \epsilon r) + O(\epsilon), \frac{\epsilon}{2} \hat{\Omega}(I_0 + \epsilon r) + O(\epsilon^2) \right)$$

and it is real-analytic on  $D^d(0, s/5) \times D^p(0, \epsilon)$ . Using (25) one has

$$\hat{\Psi}_\epsilon(r, \rho) = \left( \omega(I_0 + \epsilon r) + O(\epsilon), \frac{\epsilon}{2} (\Omega(I_0 + \epsilon r) + O(\epsilon)) \right). \quad (32)$$

Now, let  $\bar{\mu} \in \mathbb{N}_+$  and  $\bar{\beta} > 0$  denote respectively the index and the amount of nondegeneracy of  $\Psi := (\omega, \Omega)$  with respect to  $\bar{B}^d(I_0, t)$ , for some positive  $t$  independent of  $\epsilon$ . Set

$$\Psi_0(r) := (\omega(I_0 + \epsilon r), \Omega(I_0 + \epsilon r)), \quad (33)$$

$\mathcal{K}_0 := \bar{B}^d(0, s/10)$  and use definition 2 to get

$$\min_{r \in \mathcal{K}_0} \max_{0 \leq \mu \leq \bar{\mu}} \left| \partial_r^\mu |\langle c, \Psi_0(r) \rangle|^2 \right| \geq \epsilon^{\bar{\mu}} \bar{\beta} > 0$$



for every  $c \in \mathcal{S}^{d+p-1}$ .

Next, denote by  $\Psi_\epsilon$  the real-analytic function over  $D^d(0, s/5) \times D^p(0, \epsilon)$  obtained multiplying the last  $p$  component of  $\hat{\Psi}_\epsilon$  by a factor  $2/\epsilon$ . Then, observe that equations (32) and (33) imply  $\Psi_\epsilon(r, \rho) = \Psi_0(r) + O(\epsilon)$ . Therefore, denoting  $\mathcal{K}_1 := \bar{B}^p(0, \epsilon/2)$  and assuming  $\epsilon$  small enough, one has

$$\beta_\epsilon := \min_{(r, \rho) \in \mathcal{K}_0 \times \mathcal{K}_1} \max_{0 \leq \mu \leq \bar{\mu}} \left| \partial_{(r, \rho)}^\mu |\langle c, \Psi_\epsilon(r, \rho) \rangle|^2 \right| \geq \frac{\epsilon^{\bar{\mu}} \bar{\beta}}{2} > 0$$

for every  $c \in \mathcal{S}^{d+p-1}$ .

Now, if we write  $\Psi_\epsilon = (\Psi_\epsilon^{(1)}, \Psi_\epsilon^{(2)}) \in \mathbb{R}^d \times \mathbb{R}^p$ , from what observed before, it results

$$\hat{\Psi}_\epsilon = \left( \Psi_\epsilon^{(1)}, \frac{\epsilon}{2} \Psi_\epsilon^{(2)} \right).$$

Define for  $c = (c_1, c_2) \in \mathbb{R}^d \times \mathbb{R}^p$  with  $|c| = 1$  the function

$$\begin{aligned} f(r, \rho, c_1, c_2) &:= \max_{0 \leq \mu \leq \bar{\mu}} \left| \partial_{(r, \rho)}^\mu |\langle c, \Psi_\epsilon \rangle|^2 \right| \\ &= \max_{0 \leq \mu \leq \bar{\mu}} \left| \partial_{(r, \rho)}^\mu \left| \langle c_1, \Psi_\epsilon^{(1)} \rangle + \langle c_2, \Psi_\epsilon^{(2)} \rangle \right|^2 \right|; \end{aligned}$$

furthermore set

$$t_\epsilon := \sqrt{|c_1|^2 + \frac{\epsilon^2}{4} |c_2|^2}$$

and  $\bar{c}_1 = c_1 t_\epsilon^{-1}$ ,  $\bar{c}_2 = \epsilon c_2 (2t_\epsilon)^{-1}$  so that  $|(\bar{c}_1, \bar{c}_2)| = 1$ . Then one has

$$\begin{aligned} \max_{0 \leq \mu \leq \bar{\mu}} \left| \partial_{(r, \rho)}^\mu |\langle c, \hat{\Psi}_\epsilon \rangle|^2 \right| &= f\left(r, \rho, c_1, \frac{\epsilon}{2} c_2\right) = t_\epsilon^2 f\left(r, \rho, \frac{c_1}{t_\epsilon}, \frac{\epsilon}{2} \frac{c_2}{t_\epsilon}\right) \\ &\geq \frac{\epsilon^2}{4} f(r, \rho, \bar{c}_1, \bar{c}_2) \geq \frac{\epsilon^{\bar{\mu}+2} \bar{\beta}}{8} > 0 \end{aligned}$$

and it follows immediately

$$\min_{(r, \rho) \in \mathcal{K}_0 \times \mathcal{K}_1} \max_{0 \leq \mu \leq \bar{\mu}} \left| \partial_{(r, \rho)}^\mu |\langle c, \hat{\Psi}_\epsilon \rangle|^2 \right| \geq \frac{\epsilon^{\bar{\mu}+2} \bar{\beta}}{8} > 0$$

for every  $c \in \mathcal{S}^{d+p-1}$ . Since  $\mathcal{K}_0 \times \mathcal{K}_1 = \bar{B}^d(0, s/10) \times \bar{B}^p(0, \epsilon/2) = \mathcal{K}_\epsilon$  we have verified (31). In view of remark 2 we also conclude that  $\hat{\Psi}_\epsilon$  is R-nondegenerate on  $B^d(0, s/5) \times B^p(0, \epsilon)$ , provided that  $\epsilon$  is small enough.

Proposition 2 is proved. ■

### 2.3.3 Conclusion of the Proof of Theorem 4

We want to apply R  bmann's Theorem 3 to the properly degenerate case of  $\hat{H}_\epsilon$  in (16). With Propositions 1 and 2 we are in a position to meet the hypothesis of R-nondegeneracy of the frequency application required in Theorem 3. However, the ‘‘degenerate’’ case of  $\hat{H}_\epsilon$  requires that the size of its perturbation is of a sufficiently

small order in  $\epsilon$ . From (16) we see that the size of the perturbation of  $\hat{H}_\epsilon$  is order  $\epsilon^\nu$  where  $\nu$  can be chosen to be arbitrarily big<sup>24</sup> but independent of  $\epsilon$ .

Next we provide an explicit expression for the admissible size of the perturbation in Rüßmann's Theorem, i.e.  $\epsilon_0$  in (4).

**Lemma 8 (Rüßmann).** *Let  $H, \mathcal{Y}, \epsilon^*, \mathcal{A}$  and  $\tau$  be as in Theorem 3 and let  $\omega := \nabla h$  be  $R$ -nondegenerate (as in the hypotheses of Theorem 3). Consider the following quantities:*

1. *Let  $\mathcal{K} \subset \mathcal{Y}$  be any chosen compact set; let  $\mu$  be any integer greater than the index of nondegeneracy of  $\omega$  with respect to  $\mathcal{K}$  and let  $\beta$  be the “amount of nondegeneracy” corresponding to  $\mu$ .*
2. *Let  $\vartheta \in (0, 1)$  be chosen such that<sup>25</sup>  $\mathbb{T}_\vartheta^d \times (\mathcal{K} + 4\vartheta) \subset \mathcal{A}$  and define  $C_1 := |\omega|_{\mathcal{K}+3\vartheta}$ . Let  $d_0$  be the diameter of  $\mathcal{K}$ , i.e.  $d_0 := \sup_{x,y \in \mathcal{K}} |x - y|$ .*
3. *Let  $T_0 \geq e^{\frac{n+1}{\tau}}$  such that the following inequality holds*

$$\int_{T_0}^{\infty} \frac{\log T}{T^2} dT \leq \vartheta. \quad (34)$$

4. *Define*

$$C^\star := 2^{\mu+1} \frac{(\mu+1)^{\mu+2}}{\vartheta^{\mu+1}} (C_1 + 1) \quad (35)$$

*and set*

$$\gamma := (d_0^n C^\star)^{-\frac{\mu}{2}} \beta^{\frac{\mu+1}{2}} \epsilon^{\star \frac{\mu}{2}}, \quad (36)$$

$$t_0 := \frac{\gamma T_0^{-(\tau+n+1)} \vartheta}{C_1 + 1}. \quad (37)$$

5. *Finally set*

$$\begin{aligned} E_1 &:= \gamma T_0^{-(\tau+n+1)} \vartheta \\ E_2 &:= \frac{\beta t_0^\mu}{T_0 C^\star}. \end{aligned} \quad (38)$$

*Then  $\epsilon_0$  in (4) can be taken to be*

$$\epsilon_0 := c_0 \frac{\vartheta}{C_1} (\min\{E_1, E_2\})^2 \quad (39)$$

*for a suitable  $c_0 = c_0(n, \mu)$ .*

**Remark 5.** The above result follows from [16] by considering the case of maximal tori only<sup>26</sup>. More precisely:

<sup>24</sup>Recall (29) and the fact that both  $\nu_1$  and  $\nu_2$  can be arbitrarily fixed at the beginning of the process described in section 2.3.1.

<sup>25</sup>Recall definition (12)

<sup>26</sup>Compare in particular the estimates listed on page 171 of [16]; see, also, chapter 2 of [13].

- (i) The maximal case corresponds to the easier case  $p = q = 0$  in [16]. Notice, however, that it is not sufficient to substitute the values  $p = q = 0$  in R  bmann's estimates (as ,when  $p = 0$  for instance, many terms in [16, p. 171] become meaningless) but, rather, one has to go through the most of Theorem 18.5 in [16] to get the value of  $C^*$  and  $\gamma$  in (35) and (36) and through the first part of Lemma 13.4 in [16, pp. 158–161] to get the value of  $t_0$  in (37).
- (ii) (Control of small divisors) In [16, section 1.4] R  bmann introduces a so called “approximation function”  $\Phi$  in order to control the small divisors. Our choice is to take  $\Phi(T) = T^{-\tau}$  with  $\tau > n\mu$ . Comparing [16, section 1.4], one sees that such  $\Phi$  does not verify property 3, i.e.  $T^\lambda \Phi(T) \xrightarrow{T \rightarrow \infty} 0$  for any  $\lambda \geq 0$ . However, when we consider  $H = \hat{H}_\epsilon$ , we will see below that one has  $T_0 = T_{0,\epsilon} = O(\epsilon^{-2})$ ; then, equation 14.10.10 together with 13.1.4 and inequality 14.10.11 in [16] would cause  $O(\epsilon^\nu)$  to be an inadmissible size for a perturbation. Nevertheless we claim that the only decay property which is actually needed in R  bmann's Theorem 3 is

$$\lim_{T \rightarrow \infty} T^\lambda \Phi(T) = 0 \quad \text{for all} \quad 0 \leq \lambda < n\mu$$

so that our choice is perfectly suitable.

Now we are going to analyze what happens to the estimate in Lemma 8 when we consider  $\hat{H}_\epsilon$  as Hamiltonian function. In particular we are going to show that each one of the quantities appearing in Lemma 8 can be controlled by constants involving initial parameters related only to  $H_\epsilon$  in (7) and (8) times powers of  $\epsilon$ .

We point out that in the application of Theorem 3 with  $H = \hat{H}_\epsilon$  in (16) we have the following correspondences<sup>27</sup>:

$$\begin{aligned} \mathbb{T}^n &= \mathbb{T}^d \times \mathbb{T}^p, \quad x = (\theta, \zeta) \\ \mathcal{Y} &= \mathcal{Y}_\epsilon := B^d(0, s/5) \times B^p(0, \epsilon), \quad y = (r, \rho) \\ \mathcal{A} &= \mathcal{A}_\epsilon := \mathbb{T}_{\frac{s}{8}}^d \times \mathbb{T}_{\sigma_\star}^p \times D^d(0, s/5) \times D^p(0, \epsilon) \\ P &= \epsilon^\nu P_\epsilon, \quad N = N_\epsilon. \end{aligned} \tag{40}$$

Accordingly to 8.1 we consider the frequency application of the integrable part of  $\hat{H}_\epsilon$ , that is  $\hat{\Psi}_\epsilon(r, \rho)$  as in Proposition 2. We already proved that  $\hat{\Psi}_\epsilon$  is R-nondegenerate for  $(r, \rho) \in B^d(0, s/5) \times B^{2p}(0, \epsilon)$ . Now, in view of 8.1 and the correspondences in (40) we need to fix a compact set  $\mathcal{K} = \mathcal{K}_\epsilon \subset \mathcal{A}_\epsilon$ . For our convenience we take  $\mathcal{K}_\epsilon := \bar{B}^d(0, s/10) \times \bar{B}^p(0, \epsilon/2)$  so that the first inequality in (31) allows us to consider<sup>28</sup>  $\mu = \bar{\mu}$  as an integer greater than the actual index of nondegeneracy of  $\hat{\Psi}_\epsilon$  with respect to  $\mathcal{K}_\epsilon$ . Also, in view of the second inequality in (31), we can take

$$\beta = \beta_\epsilon := \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} \tag{41}$$

<sup>27</sup>See Theorem 3, Proposition 1, (12) and (21) for notations.

<sup>28</sup>Recall Proposition 2 for the definition of  $\bar{\mu}$ .

in (36) and (38).

Next, we choose  $\vartheta = \vartheta_\epsilon := \epsilon/16$  so that, for  $\epsilon$  sufficiently small and in view of (15), one has  $\mathcal{K}_\epsilon + 4\vartheta_\epsilon \subset \mathcal{A}_\epsilon$  as required in 8.2. Accordingly to Theorem 3 we also need to fix a positive number  $\epsilon^\star < \text{meas}_{d+p} \mathcal{K}_\epsilon$ . In view of our definition of  $\mathcal{K}_\epsilon$  and (15) a suitable choice is given by

$$\epsilon^\star = \epsilon^{p+1} \quad (42)$$

for  $\epsilon$  small enough.

Now, observe that the quantities  $C_1$  and  $d_0$  in Lemma 8, point 2 do not cause any change in the order in  $\epsilon$  of the size of the admissible perturbation. In fact, using equation (32) and taking  $\epsilon$  sufficiently small, we have

$$\begin{aligned} C_1 = C_{1,\epsilon} &:= |\hat{\Psi}_\epsilon|_{\mathcal{K}_\epsilon + 3\vartheta_\epsilon} \leq \\ &\leq \sup_{r \in D^d(0, s/5)} |\omega(I_0 + \epsilon r)| + \epsilon \sup_{r \in D^d(0, s/5)} |\Omega(I_0 + \epsilon r)| + O(\epsilon) \leq \\ &\leq \sup_{r \in D^d(I_0, s/5)} |\omega(r)| + \epsilon \sup_{r \in D^d(I_0, s/5)} |\Omega(r)| + O(\epsilon) \leq M_1 \end{aligned}$$

where  $M_1$  is defined in (14). Since the estimate for  $\epsilon_0$  is decreasing with respect to  $C_1$ , we can substitute  $C_1$  in (35) and (37) with  $M_1$ . The estimate for  $\epsilon_0$  is also decreasing with respect to  $d_0$  so that when we consider  $\mathcal{K} = \mathcal{K}_\epsilon$  we may simply replace  $d_0$  by 1.

Let us now analyze the quantities in 8.3 and 8.4. First of all observe that in view of (31) and  $n = d + p$  we can fix a priori an exponent  $\tau \geq (d + p)\bar{\mu}$  satisfying the requirement in 8.3. Furthermore, given the previous choice of  $\vartheta_\epsilon$ , inequality (34) becomes

$$\int_{T_0}^{\infty} \frac{\log T}{T^2} dT \leq \frac{\epsilon}{16}$$

which can be easily fulfilled, together with  $T_0 \geq e^{\frac{d+p+1}{\tau}}$ , by choosing

$$T_0 = T_{0,\epsilon} := \frac{1}{\epsilon^2} \quad (43)$$

for  $\epsilon$  sufficiently small. For what concerns the quantities defined in 8.4 we see that since  $\vartheta = \vartheta_\epsilon := \epsilon/16$  and the estimate for  $\epsilon_0$  is decreasing in  $C^\star$ , we can choose

$$C^\star = C_\epsilon^\star := 2^{5(\bar{\mu}+1)}(\bar{\mu}+1)^{\bar{\mu}+2}(M_1+1)\epsilon^{-(\bar{\mu}+1)} \quad (44)$$

having also used  $C_1 = C_{1,\epsilon} \leq M_1$ . From the fact that we can replace  $d_0$  by 1 together with equations (41), (42) and (44), one has

$$\gamma = \gamma_\epsilon := c_1(M_1+1)^{-\frac{\bar{\mu}}{2}}\epsilon^{(\bar{\mu}+1)^2 + \frac{(p+1)\bar{\mu}}{2}}\bar{\beta}^{\frac{\bar{\mu}+1}{2}}\epsilon^{\star\frac{\bar{\mu}}{2}} \quad (45)$$

for a suitable constant  $c_1 < 1$  depending only on  $\bar{\mu}$ . Moreover, given once again the previous choice of  $\vartheta = \vartheta_\epsilon$  together with equation (43) and the above definition of  $\gamma_\epsilon$ , we can replace  $t_0$  in (37) by

$$t_0 = t_{0,\epsilon} = c_1(M_1+1)^{-\frac{\bar{\mu}}{2}}\epsilon^{(\bar{\mu}+1)^2 + \frac{(p+1)\bar{\mu}}{2} + 2(\tau+d+p)+3} \quad (46)$$

for  $c_1 < 1$  as above.

From (38) we see that  $E_1$  and  $E_2$  have simple polynomial dependence on the quantities  $\gamma, T_0^{-1}, \vartheta, \beta$  and  $t_0$ . Our previous analysis shows that when we consider  $\hat{H}_\epsilon$  as Hamiltonian function, these quantities can be replaced respectively by<sup>29</sup>

$$\begin{aligned}\gamma_\epsilon &= O\left(\epsilon^{(\bar{\mu}+1)^2 + \frac{(p+1)\bar{\mu}}{2}}\right), & T_{0,\epsilon}^{-1} &= O(\epsilon^2), & \vartheta_\epsilon &= O(\epsilon) \\ \beta_\epsilon &= O(\epsilon^{\bar{\mu}+2}), & t_{0,\epsilon} &= O\left(\epsilon^{(\bar{\mu}+1)^2 + \frac{(p+1)\bar{\mu}}{2} + 2(\tau+d+p)+3}\right).\end{aligned}$$

Therefore, in view of (39) the size of the perturbation allowed by R  bmann’s Theorem when we consider  $H = \hat{H}_\epsilon$ , is order  $\epsilon^{\nu_0}$  with<sup>30</sup>

$$\nu_0 := 2\bar{\mu}^3 + (p+5)\bar{\mu}^2 + [14 + 4(\tau+d+p)]\bar{\mu} + 13. \quad (47)$$

In particular we have a condition of the form  $\epsilon_0 \leq \bar{c}\epsilon^{\nu_0}$  where  $\bar{c}$  is some positive constant independent of  $\epsilon$  and depending only on quantities related to the initial Hamiltonian  $H_\epsilon$ , namely  $\bar{\mu}, \bar{\beta}$  and  $\bar{K}$  as in Proposition 2, the Diophantine constant  $\tau \geq (d+p)\bar{\mu}$  and  $M_1$  as in (14) with  $r_0$  as in (12). By Proposition 1 we know that we can assume the size of the perturbation of  $\hat{H}_\epsilon$  to be order  $\epsilon^\nu$  for any fixed integer  $\nu \geq 4$  independent of  $\epsilon$ . Thus, by simply taking<sup>31</sup>  $\nu > \nu_0$ , we can apply R  bmann’s Theorem to  $\hat{H}_\epsilon$  and obtain Theorem 4 as a consequence. ■

### 3 Proof of Theorem 2

As it follows from the analysis described in [7, Sect. 6, pp. 1563–1569], the motions of  $(n+1)$  bodies (point masses) interacting only through gravitational attraction, restricted to the invariant symplectic submanifold of vanishing total linear momentum, are governed by the real-analytic Hamiltonian

$$F = H^0(\Lambda) + \epsilon \left( H^1(\Lambda, \xi, \eta, q, p) + H^2(\lambda, \Lambda, \xi, \eta, q, p) \right) \quad (48)$$

where:

- (i)  $(\lambda, \Lambda, \xi, \eta, q, p) \in \mathbb{T}^n \times (0, \infty)^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$  are standard symplectic coordinates;
- (ii)  $\Lambda_j = \mu_j \sqrt{M_j a_j}$ , where  $a_j > 0$  are the semi major-axis of the “instantaneous” Keplerian ellipse formed by the “Sun” (major body) and the  $j^{\text{th}}$  “planet”, while

$$\frac{1}{\epsilon \mu_j} = \frac{1}{m_0} + \frac{1}{\epsilon m_j}, \quad M_j := m_0 + \epsilon m_j,$$

$m_0$  and  $\epsilon m_j$  being, respectively the mass of the Sun and the mass of the  $j^{\text{th}}$ -planet;

<sup>29</sup>See (45), (43), (41), (46) and recall  $\vartheta = \vartheta_\epsilon := \epsilon/16$ .

<sup>30</sup>Using (47) we are able to define the values of  $\nu_1$  in (19) and (20) and  $\nu_2$  in (22) through equation (29).

<sup>31</sup>Notice that this can be done since  $\nu_0$  only depends on  $d, p, \tau$  and  $\bar{\mu}$ .

- (iii) the phase space  $\mathcal{M}$  is the open subset of  $\mathbb{T}^n \times (0, \infty)^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$  subject to the collisionless constrain

$$0 < a_n < a_{n-1} < \cdots < a_1$$

and endowed with the standard symplectic form  $\sum_{j=1}^n d\lambda_j \wedge d\Lambda_j + d\xi_j \wedge d\eta_j + dq_j \wedge dp_j$ ;

- (iv)  $H^0 := F_{\text{Kep}}$  is the Keplerian integrable limit given by

$$H^0 := F_{\text{Kep}} := \sum_{j=1}^n -\frac{\mu_j^3 M_j^2}{2\Lambda_j^2},$$

describing  $n$  decoupled two-body systems formed by the Sun and the  $j^{\text{th}}$  planet;

- (v) the “secular” Hamiltonian  $H^1$  has the form<sup>32</sup>

$$H^1 = C_0 + \sum_{j=1}^n \sigma_j \frac{\xi_j^2 + \eta_j^2}{2} + \sum_{j=1}^n \varsigma_j \frac{q_j^2 + p_j^2}{2} + O(4) \quad (49)$$

where  $C_0$ ,  $\sigma_j$  and  $\varsigma_j$  depend on  $\Lambda$ ; “ $O(4)$ ” denotes terms of order greater than or equal to four in  $(\xi, \eta, q, p)$  (and depending on  $\Lambda$ );

- (vi)  $H^2$  has vanishing average over  $\lambda \in \mathbb{T}^n$ ;  $H^i$  depend also (in a regular and non influential way) on  $\epsilon$ .

**Remark 6.** The variables  $(\lambda, \Lambda, \xi, \eta, q, p)$  are obtained from standard Poincaré variables after a rotation in  $(\xi, \eta, q, p)$  needed to diagonalize the quadratic part of the secular Hamiltonian; the “eigenvalues”  $\sigma_j$  and  $\varsigma_j$  are the the first Birkhoff invariants of the secular Hamiltonian; compare [7, pp. 1568, 1569].

The frequency map of the planetary Hamiltonian  $F$  is given by

$$\{\nu_1, \dots, \nu_n, \sigma_1, \dots, \sigma_n, \varsigma_1, \dots, \varsigma_n\}$$

where the  $\nu_j$ ’s are the Keplerian frequencies

$$\nu_j := \frac{\partial F_{\text{Kep}}}{\partial \Lambda_j} = \frac{\sqrt{M_j}}{a_j^{\frac{3}{2}}} = \frac{\mu_j^3 M_j^2}{\Lambda_j^3}. \quad (50)$$

It is customary to consider the frequency map as a function of the semi-major axes  $a$  (rather than of the actions  $\Lambda$ ); we shall therefore call the “planetary frequency map” the application<sup>33</sup>

$$\alpha : a \in \mathcal{A} \mapsto \{\nu_1, \dots, \nu_n, \sigma_1, \dots, \sigma_n, \varsigma_1, \dots, \varsigma_n\} \in \mathbb{R}^{3n} \quad (51)$$

<sup>32</sup>There is a difference of a factor  $\frac{1}{2}$  with the notations used in Ref. [7]. The computations are performed in [11].

<sup>33</sup>Obviously, the property of being R-nondegenerate can be equivalently discussed in terms of the  $\Lambda$ ’s or in terms of the  $a$ ’s.

where

$$\mathcal{A} := \{(a_1, a_2, \dots, a_n) \in \mathbb{R}^n : 0 < a_n < a_{n-1} < \dots < a_1\}.$$

Clearly, the idea is to apply Theorem 4 to the real-analytic Hamiltonian  $F$  in (48) with  $d = n$  and  $p = 2n$ :  $(\varphi, I)$  corresponding to  $(\lambda, \Lambda)$  here and  $u$  corresponding to  $(\xi, q)$  and  $v$  to  $(\eta, p)$ . However, it turns out that *the main hypothesis of Theorem 4 does not hold*, namely, the planetary frequency map  $\alpha$  is  $\mathbb{R}$ -degenerate: in fact (up to rearranging the  $(q, p)$ -variables) one has

$$\begin{cases} \varsigma_n = 0, \\ \sum_{j=1}^n (\sigma_j + \varsigma_j) = 0. \end{cases} \quad (52)$$

The first relation is related to the rotation invariance of the system; the second relation seems to have been noticed (at least in this generality) for the first time by Michael Herman and is therefore normally referred to as the “Herman resonance”.

The two resonances in (52) are, however, the only linear relations identically satisfied; in fact in [7, Proposition 78, p. 1575] it is proved the following

**Proposition 3.** *For all  $n \geq 2$  there exists an open and dense set with full Lebesgue measure  $U \subset \mathcal{A}$ , where  $\alpha_j \neq \alpha_i$  whenever  $j \neq i$  and the following property holds: for any open and simply connected set  $V \subset U$ , the  $\alpha_j$  define  $3n$  holomorphic functions and if*

$$\alpha \cdot (c^1, c^2, c^3) = \nu \cdot c^1 + \sigma \cdot c^2 + \varsigma \cdot c^3 \equiv 0$$

for some  $c^i \in \mathbb{R}^n$ , then

$$\begin{cases} \text{either } c^1 = 0, c^2 = 0, c^3 = (0, \dots, 0, 1), \\ \text{or } c^1 = 0, c^2 = (1, \dots, 1) = c^3. \end{cases} \quad (53)$$

In order to remove the secular resonances (52), we consider the following “extended Hamiltonian” on

$$\tilde{\mathcal{M}} := \mathcal{M} \times \mathbb{T} \times \mathbb{R}$$

adding a pair of conjugate symplectic variables<sup>34</sup>  $(\theta_\rho, \rho) \in \mathbb{T} \times \mathbb{R}$ :

$$\tilde{F} := F + \frac{\rho^2}{2} + \epsilon \rho^2 C_z \quad \text{with} \quad C_z := \sum_{j=1}^n \left( \Lambda_j - \frac{1}{2} (\xi_j^2 + \eta_j^2 + q_j^2 + p_j^2) \right). \quad (54)$$

Let us make a few comments.

- (vii)  $C_z$  is the vertical component of the total angular momentum in Poincaré variables (compare [12] and also formula (44) in [7]); the form of  $C_z$  is unchanged in the above variables  $(\xi, \eta, q, p)$ , which are obtained from the Poincaré variables by an orthogonal transformation.

---

<sup>34</sup>I.e.,  $\tilde{\mathcal{M}}$  is endowed with the symplectic form  $\sum_{j=1}^n (d\lambda_j \wedge d\Lambda_j + d\xi_j \wedge d\eta_j + dq_j \wedge dp_j) + d\theta_\rho \wedge d\rho$ .

- (viii) Since  $C_z$  is an integral for  $F$  (i.e., Poisson commutes with  $F$ ),  $F$  and  $\tilde{F}$  Poisson commutes:

$$\{F, \tilde{F}\} = \{\tilde{F}, F\} = 0$$

where  $\{\cdot, \cdot\}$  and  $\{\cdot, \cdot\}$  denote, respectively, the Poisson bracket on  $\tilde{\mathcal{M}}$  and on  $\mathcal{M}$ ; clearly, since  $\tilde{F}$  does not depend explicitly upon the angle  $\theta_\rho$ , also  $\rho$  is an integral for  $\tilde{F}$  (and for  $F$ ).

This fact will be important later since from Lagrangian intersection theory it follows that two commuting Hamiltonians have, in general, the same Lagrangian tori (see item (x) below for the precise statement).

- (ix) The extended Hamiltonian  $\tilde{F}$  may be rewritten as

$$\tilde{F} = \tilde{H}^0 + \epsilon(\tilde{H}^1 + H^2)$$

with

$$\begin{aligned} \tilde{H}^0 &:= F_{\text{Kep}}(\Lambda) + \frac{\rho^2}{2}, \\ \tilde{H}^1 &:= C_0(\Lambda) + \rho^2 \sum_{j=1}^n \Lambda_j + \\ &\quad + \sum_{j=1}^n (\sigma_j - \rho^2) \frac{\xi_j^2 + \eta_j^2}{2} + \sum_{j=1}^n (\varsigma_j - \rho^2) \frac{p_j^2 + q_j^2}{2} + O(4). \end{aligned} \quad (55)$$

Thus, the “slow” action variables<sup>35</sup> are  $I = (\rho, \Lambda_1, \dots, \Lambda_n)$  and the (extended) planetary frequency map is given by

$$\tilde{\alpha} : (\rho, a) \in \mathcal{A} \times \mathbb{R} \mapsto \tilde{\alpha}(\rho, a) := ((\rho, \nu), \tilde{\sigma}, \tilde{\varsigma}) \in \mathbb{R}^{n+1} \times \mathbb{R}^n \times \mathbb{R}^n$$

with

$$\tilde{\sigma}_j := \sigma_j - \rho^2, \quad \tilde{\varsigma}_j := \varsigma_j - \rho^2.$$

Proposition 3, implies immediately that  $\tilde{\alpha}$  is  $\mathbb{R}$ -nondegenerate: suppose, in fact, that

$$\tilde{\alpha} \cdot ((c, c^1), c^2, c^3) = \rho c + \nu \cdot c^1 + \tilde{\sigma} \cdot c^2 + \tilde{\varsigma} \cdot c^3 \equiv 0$$

for some  $c \in \mathbb{R}$  and  $c^i \in \mathbb{R}^n$ ; such expression is a second order polynomial in  $\rho$  and in order to vanish identically have to vanish its coefficients, i.e.,

$$\nu \cdot c^1 + \sigma \cdot c^2 + \varsigma \cdot c^3 = 0, \quad (56)$$

$$c = 0,$$

$$-\sum_{j=1}^n c_j^2 + c_j^3 = 0. \quad (57)$$

But then, by Proposition 3 (and because of (56)), one must have one of the alternatives listed in (53), which are incompatible with (57).

Thus  $\tilde{\alpha}$  is  $\mathbb{R}$ -nondegenerate as claimed and Theorem 4 can be applied to the extended Hamiltonian<sup>36</sup>  $\tilde{F}$ , yielding, for  $\epsilon$  small enough, a positive measure

<sup>35</sup>Corresponding in Theorem 4 to  $I = (I_1, \dots, I_d)$ ,  $d = n + 1$ ; compare also footnote 36 below.

<sup>36</sup>The correspondence with the notation of Theorem 4 being:  $d = n + 1$ ,  $p = 2n$ ,  $H_\epsilon = \tilde{F}$ ,  $f = \tilde{H}^1 + H^2$ ,  $I = (\Lambda, \rho)$ ,  $(u, v) = ((\xi, q), (\eta, p))$ ,  $h(I) = H^0$ ,  $f_{00} = C_0 + \rho^2 \sum_{j=1}^n \Lambda_j$ ,  $\omega = (\nu, \rho)$ ,  $\Omega = (\tilde{\sigma}, \tilde{\varsigma})$ .



set of real-analytic  $(3n + 1)$ -dimensional Lagrangian tori in  $\tilde{\mathcal{M}}$  invariant for  $\tilde{F}$  and carrying quasi-periodic motion with Diophantine frequencies.

The fact that  $\tilde{F}$  is independent of  $\theta_\rho$  and that  $\omega_* := \partial_\rho \tilde{F} = (\rho + 4\epsilon\rho C_z)$  is constant along  $\tilde{F}$ -trajectories (compare point (viii) above) implies immediately that the tori  $\mathcal{T} \subset \tilde{\mathcal{M}}$  obtained through Theorem 4 have the following parametrization

$$\mathcal{T} := \left\{ (Z(\psi, \theta_\rho), \theta_\rho, \rho) : (\psi, \theta_\rho) \in \mathbb{T}^{3n} \times \mathbb{T} \right\} \quad (58)$$

where  $Z \in \mathcal{M}$  and with  $\tilde{F}$ -flow given by

$$\phi_{\tilde{F}}^t(Z(\psi, \theta_\rho), \theta_\rho, \rho) = (Z(\psi + \omega t, \theta_\rho + \omega_* t), \theta_\rho + \omega_* t, \rho),$$

for a suitable vector  $\omega \in \mathbb{R}^{3n}$ , so that  $(\omega, \omega_*)$  forms a Diophantine vector in  $\mathbb{R}^{3n+1}$ .

- (x) In [7, Lemma 82, p. 1578] the following statement is proved

*If  $F$  and  $G$  are two commuting Hamiltonians and if  $\mathcal{T}$  is a Lagrangian torus invariant for  $F$  and with a dense  $F$ -orbit, then it is also  $G$ -invariant.*

Thus, since  $\tilde{F}$  and  $F$  (viewed as functions on  $\tilde{\mathcal{M}}$ ) commute, the tori obtained in (ix) (on which any  $\tilde{F}$ -orbit is dense) are also invariant for the flow on  $\tilde{\mathcal{M}}$  generated by  $F$ . Furthermore, the  $F$ -flow in  $\tilde{\mathcal{M}}$  leaves both  $\theta_\rho$  and  $\rho$  fixed so that, for any fixed  $\theta_\rho \in \mathbb{T}$ , the  $3n$ -dimensional torus

$$\mathcal{T}_{\theta_\rho} := \left\{ (Z(\psi, \theta_\rho), \theta_\rho, \rho) : \psi \in \mathbb{T}^{3n} \right\}$$

is invariant for  $F$ . But this means that such tori are invariant also for the  $F$ -flow in  $\mathcal{M}$ , finishing the proof of Theorem 2.  $\blacksquare$

**Remark 7.** The strategy followed here is similar to that followed in [7] with a few differences: first, in [7]  $\rho$  is treated as a dumb parameter and no extended phase space is introduced (but an extra argument is then needed to discuss the nondegeneracy of the frequency map with respect to parameters and to discuss the measure of the tori obtained); secondly, in [7] there is a restriction to a fixed vertical angular momentum submanifold, which is not needed here.

## References

- [1] Arnol'd, V.I. Small Denominators and Problems of Stability of Motion in Classical and Celestial Mechanics. *Uspehi Mat. Nauk*, 18(6(114)):91–192, 1963.
- [2] Arnol'd, V.I., Kozlov, V.V. and Neishtadt A.I. (editors). *Mathematical Aspects of Classical and Celestial Mechanics*, volume 3 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, 3rd edition, 2006.
- [3] Biasco, L., Chierchia, L. and Valdinoci, E. Elliptic Two-Dimensional Invariant Tori for the Planetary Three-Body Problem. *Arch. Rational Mech. Anal.*, 170:91–135, March 2003.
- [4] Broer, H. and Takens, F. Unicity of Kam Tori. *Ergod. Th. & Dynam. Sys.*, 27:713–724, 2007.

- [5] Celletti, A. and Chierchia, L. Kam stability and celestial mechanics. *Memoirs of the AMS*, 187(878):134 pp., 2007.
- [6] Fathi, A., Giuliani, A. and Sorrentino, A. Uniqueness of Invariant Lagrangian Graphs in a Homology or a Cohomology class, 2007. Preprint.
- [7] Féjoz, J. Démonstration du ‘théorème d’Arnold’ sur la stabilité du système planétaire (d’après M. Herman). *Ergod. Th. & Dynam. Sys.*, 24(5):1521–1582, 2004.
- [8] Féjoz, J. Version révisée de l’article paru dans le Michael Herman Memorial Issue Ergodic Theory Dyn. Sys. 24(5) : 1521–1582 (2004), 2007. <http://www.institut.math.jussieu.fr/~fejoz/arnold.pdf>  
See also the *erratum* at <http://hal.archives-ouvertes.fr/ccsd-00079059>.
- [9] Herman, M.R. Démonstration d’un théorème de V.I. Arnold. *Séminaire de Systèmes Dynamiques et manuscrits*, 1998.
- [10] Hofer, H. and Zehnder, E. *Symplectic Invariants and Hamiltonian Dynamics*. Birkhäuser Verlag, 1994.
- [11] Laskar, J. Analytical framework in Poincaré variables for the motion of the solar system. *Predictability, Stability and Chaos in n-Body Dynamical Systems*, pages 93–114, 1991.
- [12] Poincaré, H. *Leçons de Mécanique Céleste*, 1905–1907. Gauthier-Villars, Paris.
- [13] Pusateri, F. Analytic KAM Tori for the planetary (n+1)-body problem. Master’s thesis, Università degli Studi Roma Tre, October 2006. [www.mat.uniroma3.it/users/chierchia/TESI](http://www.mat.uniroma3.it/users/chierchia/TESI).
- [14] Pyartli, A.S. Diophantine approximations on Euclidean submanifolds. *Functional. Anal. Appl.*, 3:303–306, 1969. (in Russian) *Funcckional. Anal. i Prilozhen*, volume 3, pages 59–62.
- [15] Robutel, P. Stability of the Planetary Three-body Problem. *Celestial Mech. Dynam. Astronom.*, 62(3):219–261, 1995. II. KAM theory and existence of quasi-periodic motions.
- [16] Rüßmann, H. Invariant Tori in Non-Degenerate Nearly Integrable Hamiltonian Systems. *R. & C. Dynamics*, 2(6):119–203, March 2001.
- [17] Salamon, D. The Kolmogorov-Arnold-Moser theorem. *Math. Phys. Electron. J.*, 3:37, 2004. (electronic).
- [18] Sevryuk, M.B. The classical kam theory and the dawn of the twenty-first century. *MOSCOW MATHEMATICAL JOURNAL*, 3(3):1113–1144, July-September 2003.